

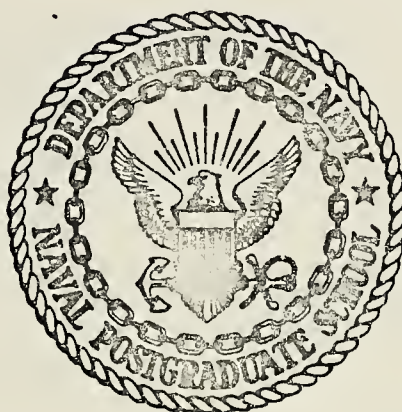
LANGUAGE IDENTIFICATION BY STATISTICAL ANALYSIS

Morton David Rau

DUDLEY KNOX LIBRARY
NAVAL POSTGRADUATE SCHOOL
MONTEREY, CALIFORNIA 93940

NAVAL POSTGRADUATE SCHOOL

Monterey, California



THESIS

LANGUAGE IDENTIFICATION BY STATISTICAL ANALYSIS

by

Morton David Rau

September 1974

Thesis Advisors:

R. A. Weitzman, V. M. Powers

Approved for public release; distribution unlimited.

T 164081

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) Language Identification by Statistical Analysis		5. TYPE OF REPORT & PERIOD COVERED Master's Thesis September 1974
7. AUTHOR(s) Morton David Rau		6. PERFORMING ORG. REPORT NUMBER
9. PERFORMING ORGANIZATION NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		8. CONTRACT OR GRANT NUMBER(s)
11. CONTROLLING OFFICE NAME AND ADDRESS Naval Postgraduate School Monterey, California 93940		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS
14. MONITORING AGENCY NAME & ADDRESS (if different from Controlling Office) Naval Postgraduate School Monterey, California 93940		12. REPORT DATE September 1974
		13. NUMBER OF PAGES 128
		15. SECURITY CLASS. (of this report) Unclassified
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE
16. DISTRIBUTION STATEMENT (of this Report) Approved for public release; distribution unlimited.		
17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)		
18. SUPPLEMENTARY NOTES		
19. KEY WORDS (Continue on reverse side if necessary and identify by block number) Language Identification, Linguistic Analysis, Linguistics, Statistical Linguistics		
20. ABSTRACT (Continue on reverse side if necessary and identify by block number) An analysis was conducted of English and Spanish text. The statistical analysis determined the independent probability of letters and the joint probability of various letter combinations for large samples of each language. Various methods were tested in an attempt to utilize these characteristics to identify the language of a short sample text. By use of the joint		

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

probability of various vowel-consonant relationships and the Kolmogorov-Smirnov Goodness of Fit Test an identification system was defined that provided a significance level of .0077 for a sample of 107 letters (approximately 21 words). Investigation also showed that the space rate or the interword structure in each language contains a measure of intelligence and was useful in identification.

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE(When Data Entered)

Language Identification

by

Statistical Analysis

by

Morton David Rau
Lieutenant Commander, United States Navy
B.S., Tulane University, 1962

Submitted in partial fulfillment of the
requirements for the degree of

MASTER OF SCIENCE IN MANAGEMENT

from the

NAVAL POSTGRADUATE SCHOOL
September 1974

Thes 5
R 2.40
31

ABSTRACT

An analysis was conducted of English and Spanish text. The statistical analysis determined the independent probability of letters and the joint probability of various letter combinations for large samples of each language.

Various methods were tested in an attempt to utilize these characteristics to identify the language of a short sample text. By use of the joint probability of various vowel-consonant relationships and the Kolmogorov-Smirnov Goodness of Fit Test an identification system was defined that provided a significance level of .0077 for a sample of 107 letters (approximately 21 words). Investigation also showed that the space rate or the interword structure in each language contains a measure of intelligence and was useful in identification.

TABLE OF CONTENTS

I.	INTRODUCTION -----	8
II.	BACKGROUND -----	10
	A. ANALYTICAL CONSIDERATIONS OF THE MODERN LATIN ALPHABET -----	10
	B. YULE'S CHARACTERISTIC (K) -----	11
	C. KOLMOGOROV-SMIRNOV TEST OF GOODNESS OF FIT -----	13
III.	EXPERIMENTS -----	15
	A. GENERAL ANALYSIS CONSIDERATIONS -----	15
	B. GENERAL PROGRAMING CONSIDERATIONS -----	15
	C. POPULATION PARAMETERS -----	16
	D. TESTING PROCEDURES -----	18
	1. Testing by Yule's K -----	19
	2. Testing by Kolmogorov-Smirnov Goodness of Fit Test -----	20
	E. RESULTS -----	21
IV.	CONCLUSIONS -----	24
APPENDIX A.	English Population Values (V-1) -----	26
APPENDIX B.	Spanish Population Values (V-1) -----	32
APPENDIX C.	English Population Values (V-2) -----	38
APPENDIX D.	Spanish Population Values (V-2) -----	44
APPENDIX E.	Significance Levels for Kolmogorov-Smirnov Test -----	50
APPENDIX F.	Final Test Sources -----	51
APPENDIX G.	Population Analysis Program (V-1) -----	57
APPENDIX H.	Population Analysis Program (V-2) -----	71

APPENDIX I. Language Identification Program (V-1) -----	87
APPENDIX J. Language Identification Program (V-2) -----	106
LIST OF REFERENCES -----	126
INITIAL DISTRIBUTION LIST -----	128

LIST OF TABLES

Table	Title	Page
I.	Proportion of Correct Identifications -----	22
II.	English Joint Probability Table for Vowel-Consonant -----	26
III.	English (V-1) Independent Probability of X_n -----	27
IV.	English (V-1) Joint Probability Table for X_n, X_{n+1} -----	28
V.	English (V-1) Joint Probability Table for $X_n, -, X_{n+2}$ -----	30
VI.	Spanish Joint Probability Table for Vowel-Consonant -----	32
VII.	Spanish (V-1) Independent Probability of X_n -----	33
VIII.	Spanish (V-1) Joint Probability Table for X_n, X_{n+1} -----	34
IX.	Spanish (V-1) Joint Probability Table for $X_n, -, X_{n+2}$ -----	36
X.	English (V-2) Independent Probability of X_n -----	39
XI.	English (V-2) Joint Probability Table for X_n, X_{n+1} -----	40
XII.	English (V-2) Joint Probability Table for $X_n, -, X_{n+2}$ -----	42
XIII.	Spanish (V-2) Independent Probability of X_n -----	45
XIV.	Spanish (V-2) Joint Probability Table for X_n, X_{n+1} -----	46
XV.	Spanish (V-2) Joint Probability Table for $X_n, -, X_{n+2}$ -----	48
XVI.	Significance Levels for Kolmogorov-Smirnov Test -----	50

I. INTRODUCTION

In today's world of ever increasing written collections it requires a certain level of expertise to properly identify and file the material according to its language. In addition to libraries this requirement exists within large government agencies such as the State Department and in various intelligence collection agencies. Although it would be expedient for the staff of these repositories to recognize and understand every language used it does not seem to be a practical requirement. It would at least seem desirable that these persons be provided with a simple method for identifying the language even if they were unable to read the words. An approach to this problem would be to have the clerk utilize a computer terminal to enter a short sample of the text which would be analyzed by the computer and then display the identified language on the terminal. The question that arises is how would the computer determine what language the sample represented. This was the problem approached in this research.

There have been many works published on the subject of mathematical linguistics most of which seem to concern themselves with identifying the author of various literary works or style analysis of individuals or regions. One investigator, Yukio Nakamura, has developed a language identification method for some twenty-five languages based upon the probability of occurrence of various specific characters and words in each language [Ref. 1].

Since language is basically a code used to convey information it would seem that each should have a particular characteristic that would

allow identification based on a statistical pattern. It was the intent of this research to design a set of computer programs which would analyze character frequency and order of occurrence in languages which are written in the Modern Latin Alphabet and then to utilize this data to identify an unknown sample.

II. BACKGROUND

A. ANALYTICAL CONSIDERATIONS OF THE MODERN LATIN ALPHABET

The Modern Latin Alphabet consists of the basic twenty-six characters, additional special characters and characters combined with diacritical signs. Very few languages utilize only the basic twenty-six characters. English is one which relies upon the basic twenty-six, except for loan words and less fashionable use of diacritics, such as 'coördinate'.

Many languages utilize additional characters which can be divided into three classifications:

1. Single characters, e.g., ß in German.

2. Joint characters, e.g., œ in French.

3. Combined characters, e.g., ch in German or Spanish. In some languages these combinations are looked upon as individual letters and have their appropriate place in the alphabet. For the purposes of this analysis they have been considered to be two individual letters.

Diacritical marks are generally used to indicate a difference in pronunciation of a letter but in some languages the marked letter is considered to be a distinct letter, such as the ü in German. In basic linguistics they are classified as modified letters and accordingly for the purposes of this study were treated as forms of the basic letter.

Having defined the criteria for the alphabet to be analyzed the question which arose was to what extent should the interrelationship of letters be considered. It was decided that limiting the interrelationship considerations to three adjacent letters would provide sufficient

indication of trends in letter relationships without broaching too greatly into the influences of word structure. Given any three adjacent characters (X_n, X_{n+1}, X_{n+2}) it was decided to evaluate the frequencies of the following events:

1. Independent occurrence of each character, X_n .
2. Joint occurrence of each pair, X_n, X_{n+1} .
3. $X_n, -, X_{n+2}$: the joint occurrence of a given letter X_{n+2} as the second letter following a given X_n .
4. Vowel-Consonant relationship in the form of the occurrence of:
 - a. vowel following vowel;
 - b. consonant following vowel;
 - c. vowel following consonant;
 - d. consonant following consonant.

Such characteristics as the occurrence of X_n, X_{n+1} generate 676 (26×26) combinations for the English Language. This has to be compared to at least the same number from any given sample. An extensive review of works on statistical linguistic analysis and nonparametric statistics resulted in the selection of two tests as a possible means of comparing a sample with a language population. These were a characteristic 'K' devised by the British statistician G. Udny Yule and the Kolmogorov-Smirnov Test of Goodness of Fit, which are discussed below.

B. YULE'S CHARACTERISTIC (K)

G. Udny Yule was involved in conducting statistical analysis of accident distribution among various people when faced with the problem of measuring the liability of classes of people independent of the period of exposure to risk or the total number of accidents met by the whole

group at risk. As a solution to this problem he developed an equation which yielded what he termed the 'characteristic' which was nothing more than a measure of the whole group's liability. By comparing accident rate or liability to repetitiveness of words in a literary work he proposed the use of this characteristic as a measure of repetitiveness in literature. The characteristic is expressed simply as a numerical value, such as 10.6, or 97.5, or 874.3. Since the analysis of each language and sample was to measure frequency of occurrence it was felt that this measure could be said to be analogous to repetitiveness as defined by Yule.

In general, Yule's approach was to determine for a given literary work how many common nouns occurred (were used) once, how many occurred twice, thrice, etc. He then performed the following calculations to obtain the characteristic (K).

X = occurrence category, i.e., X times occurring

f_x = frequency of X , i.e., the number of words occurring X times

$$S_1 = \sum_{x=1}^n f_x X$$

$$S_2 = \sum_{x=1}^n f_x X^2$$

$$K = 10,000 \frac{S_2 - S_1}{(S_1)^2}$$

The number 10,000 was introduced simply to avoid the inconvenience of handling very small numbers [Ref. 2].

Adaptation of this measurement to the $m \times n$ matrix generated by determination of letter occurrences was as follows:

$$S_1 = \sum_{x=1}^n f_x X = f_1(1) + f_2(2) + \dots + f_n(n)$$

$$S_2 = \sum_{x=1}^n f_x X^2 = f_1(1)^2 + f_2(2)^2 + \dots + f_n(n)^2$$

Instead of X being used to relate the frequency of a word it was equated to the frequency of a specific letter combination.

∴ Let: $X = Y_{m,n}$ = frequency of occurrence of the mn letter combination

$$f_x = f_{Y_{m,n}} = 1$$

$$\therefore S_1 = 1(Y_{1,1}) + 1(Y_{1,2}) + \dots + 1(Y_{m,n})$$

$$S_1 = \sum_{m=1}^m \sum_{n=1}^n Y_{m,n}$$

Similarly:

$$S_2 = \sum_{m=1}^m \sum_{n=1}^n (Y_{m,n})^2$$

This adaptation yields a single value or measure of repetitiveness in a specific category for each language population and sample.

C. KOLMOGOROV-SMIRNOV TEST OF GOODNESS OF FIT

Let X be a random variable with the continuous probability distribution function $f(x) = \text{Prob} [X \leq x]$ and let X_1, X_2, \dots, X_n be a sample of independent variables of size n for X , ordered so that $X_1 \leq X_2 \leq \dots \leq X_n$ which determines the empirical distribution function

$$f_n(x) = \begin{cases} 0 & \text{for } x < X_1 \\ (k/n) & \text{for } X_k \leq x \leq X_{k+1} \quad k = 1, 2, \dots, n-1, \\ 1 & \text{for } X_n \leq x \end{cases}$$

It should be expected that for n large, $f_n(x)$ would very likely approach $f(x)$. Kolmogorov defined the statistic

$$D_n = \text{least upper bound of } \left| f(x) - f_n(x) \right|$$

which measures the greatest absolute difference between $f(x)$ and $f_n(x)$.

He further showed that it has the following properties which assists in determining how closely $f_n(x)$ represents the population $f(x)$:

1. the probability distribution of D_n depends on n but is independent of $f(x)$, (the maximum D_n is a random variable)

2. the probability distribution of D_n is given by the relationship

$$\lim_{n \rightarrow \infty} \text{Prob} \left[D_n < \frac{z}{\sqrt{n}} \right] = 1 - 2 \sum_{k=1}^{\infty} (-1)^{k-1} e^{-2k^2 z^2} = L(z)$$

[Refs. 3, 4, 5, and 6].

The function $L(z)$ has been tabulated by N. Smirnov [Ref. 7]. From the equation $D_n \sqrt{n} = z$ a value of z is obtained to enter the tables and thereby determine $\text{Prob} \left[D_n < \frac{z}{\sqrt{n}} \right]$ or find the significance level of the fit as $1 - L(z)$. These values have been computed and are tabulated in Appendix E.

Prior to continuing some comment should be made as to the selection of the Kolmogorov-Smirnov Test over the Chi-Squared Test. The Chi-Squared test requires a large sample size, because n at finite values is not distribution-free and only approaches this condition as n approaches infinity [Refs. 4 and 8].

III. EXPERIMENTS

A. GENERAL ANALYSIS CONSIDERATIONS

In considering the alphabetic structure of a language there is no doubt as to the fact that the letters carry a certain amount of intelligence but it was not certain as to what effect if any was played by the space, other than to separate words. In order to determine the effect of the space as a character in identifying a language two analytical approaches were taken in evaluating the characteristics described in section II. A. The first (version one, Appendix G) was to conduct the analysis considering spaces as separators, not characters. All characters not described as alphabetic, such as the period, comma, etc., were considered to be spaces and signified only the termination of a word. Joint frequencies such as that of X_n, X_{n+1} were counted only when the pair occurred within the same word. This resulted in the measurement of the desired characteristics exclusively within words, without measuring the interword relationships. The second approach (version two, Appendix H) was to consider the space and nonalphabetic characters between words as a space character. This resulted in the measurement of such characteristics as frequency of first letter occurrence, frequency of last letter occurrences, and the interword relationship of last letter to first letter.

B. GENERAL PROGRAMING CONSIDERATIONS

All computer programs were written in American National Standard COBOL.

Input language samples were fed to the programs via standard eighty column punch cards in the following format:

<u>Card Column</u>	<u>Content</u>
1, 2	Blank
3 - 72	Text
73	Blank unless the last word on the previous card could not be completed in column 72, then a '-' was placed in column 73 to signify continuation on the next card.
74 - 80	Sequence Number

Only alphabetic text with normal punctuation was used for input text. Numeric characters appearing in the various text material were converted to their alphabetic equivalent. To allow for the compilation of data from languages which utilize special characters the keypunch characters 1, 2, 3, 4, and 5 were allocated to those classified as vowels while 6 and 7 were for consonants. This conversion, if required, had to be performed by the operator at the time of entry or during data preparation.

In those programs that were written to consider the space as a character of the alphabet, all interword spaces and punctuation marks were compressed to one space.

C. POPULATION PARAMETERS

The actual population in the case of written text is potentially infinite. Due to limited resources it was decided to test the theory using representative samples of two languages as a measure of their respective populations. The source of these population samples was

chosen to be Reader's Digest as it is available in many languages and contains several short articles, by different authors, which have been selected from various areas of interest, thereby suppressing the influence of one person's or region's style on the assumed population.

The English population consisted of text material, excluding advertisements, from Reader's Digest, Vol. 104, No. 623, March 1974, pages 8, 11, 12, 31 - 90, 109 and 110. This text consists of 79,185 characters excluding spaces and 96,100 characters including spaces.

The Spanish population consisted of text material, excluding advertisements, from Selecciones Del Reader's Digest, Vol. 67, No. 402, May 1974, pages 3 - 10, 21 - 30, 37-60, and 75 -80. This text consists of 78,851 characters excluding spaces and 94,992 characters including spaces.

Both populations were analyzed for the characteristics described in section II. A. (with and without space consideration) and the following population means were obtained:

1. Yule's 'K' based on the independent occurrence of each character, X_n ;
2. Yule's 'K' based on the joint occurrence of X_n, X_{n+1} ;
3. Yule's 'K' based on the joint occurrence of $X_n, -, X_{n+2}$;
4. Probability table for the independent occurrence of X_n ;
5. Joint probability table for the occurrence of X_n, X_{n+1} ;
6. Joint probability table for the occurrence of $X_n, -, X_{n+2}$;
7. Joint probability table for vowel-consonant occurrence.

(A tabulation of these values is contained in Appendices A, B, C, and D)
Each probability table was formed into a cumulative relative frequency

curve, the points of which were transferred to punched cards, for later use in testing samples by the Kolmogorov-Smirnov Goodness of Fit Test.

In order to facilitate the use of Yule's 'K' in testing small samples it was necessary to approximate the standard deviation (σ) for each value of Yule's 'K'. The population was divided into groups of ten card samples from which a new 'K' was determined for each sample. The following calculations were then performed:

K = Yule's 'K' obtained from each sample of 10 cards (less than 700 characters).

n = number of 10 card samples contained in the population.

$\bar{K} = \frac{\sum K}{n}$ = mean value of Yule's 'K'.

$$\therefore \sigma_{10} = \sqrt{\frac{(\sum K^2 - n\bar{K})}{(n - 1)}}$$

σ_{10} was then adjusted to a population standard deviation based on one card samples by computing

$$\sigma = \sigma_{10} \sqrt{10}$$

[Ref. 8].

(Values of σ are contained in Appendices A, B, C, and D.) A histogram was plotted for the various values of Yule's 'K' and they appeared to conform to a normal distribution.

D. TESTING PROCEDURES

In constructing testing procedures there were two approaches which were considered. The first was that a small sample would be analyzed and identified as either English or Spanish. The second approach was to analyze a small sample and then report it as either English, Spanish, or undecisive based upon some predefined constraints. Since it was

intended, if the theory worked, that the process could be expanded to cover all languages written in the Modern Latin Alphabet it was felt that the first approach was more realistic (useful).

1. Testing by Yule's K

Since the incremental population samples of ten cards appeared to conform to the normal distribution it was assumed that for all languages the value of K obtained from small samples would conform to the normal distribution. For each sample tested three values of K were obtained, one for each of the three K tests. Each value of K was then compared to its respective population mean, μ , obtained from the representative population analysis, in each language in order to determine the appropriate normal deviate, z, by

$$z = \frac{K - \mu}{\sigma / \sqrt{n}}$$

where n was the number of data cards in that particular sample. For each of the three tests the z for English was compared with the z for that test in Spanish. The smallest absolute value of z identified the language or that language to which the sample had the highest probability of belonging. The smallest z from each test was then used to compute the significance level, α , as a measure of how closely K and μ matched based on the probability of a normal distribution. For absolute values of z less than 4.0, α was computed as follows:

$$\begin{aligned} x &= |z| \\ \alpha &= 1 - 2x(.39894,22804,014) \left[1 - \frac{1}{6}x^2 + \frac{P}{6.66666,66666,67} x^2 \right. \\ &\quad \left. - \frac{P}{8.4} + \frac{P}{10.28571,42857,1} x^2 - \frac{P}{12.22222,22222,2} x^2 \right. \\ &\quad \left. + \frac{P}{14.18181,81818,2} x^2 - \frac{P}{16.15384,61538,5} x^2 \right] \end{aligned}$$

$$\begin{aligned}
& + \frac{P}{18.13333,33333,3} x^2 - \frac{P}{20.11764,70588,2} x^2 \\
& + \frac{P}{22.10526,31578,9} x^2 - \frac{P}{24.09523,80952,4} x^2 \\
& + \frac{P}{26.08695,65217,4} x^2 - \frac{P}{28.08000,0000,0} x^2 \\
& + \frac{P}{30.07407,40740,7} x^2 - \frac{P}{32.06896,55172,4} x^2 \\
& + \frac{P}{34.06451,61290,3} x^2 - \frac{P}{36.06060,60606,1} x^2 \\
& + \frac{P}{38.05714,28571,4} x^2 - \frac{P}{40.05405,40540,5} x^2 \\
& + \frac{P}{42.05128,20512,8} x^2 - \frac{P}{44.04878,04878,0} x^2 \\
& + \frac{P}{46.04651,16279,1} x^2 - \frac{P}{48.04444,44444,4} x^2 \\
& + \frac{P}{50.04255,3191} x^2 - \frac{P}{52.04081,6326} x^2 \\
& + \frac{P}{54.03921,5686} x^2 - \frac{P}{56.03773,5849} x^2 \\
& + \left[\frac{P}{58.03636} x^2 - \frac{P}{60.03509} x^2 + \frac{P}{62.03} x^2 \right]
\end{aligned}$$

where P = the absolute value of the preceding term [Ref. 9]. For absolute values of z greater than or equal to 4.0, α was reported as being less than 0.00006 [Ref. 8].

2. Testing by the Kolmogorov-Smirnov Goodness of Fit Test

For each sample tested four cumulative relative frequency tables were computed, one from each of the probability tables described in section III. C. These tables represented points, $f_n(x)$, on the cumulative relative frequency (CRF) curve for each particular sample. The CRF curve points for the sample were compared against the estimate of the respective curve points, $f(x)$, along the ordinate, for the Spanish and English populations. From each comparison the value of

D_n = least upper bound of $|f(x) - f_n(x)|$ was obtained. For each of the four tests the value of D_n from the comparison with Spanish was compared to that D_n from the comparison with English. The sample was judged to be of that language which resulted in the smallest D_n . The smallest value of D_n from each test was then used to compute the significance level as a measure of how closely the sample and population CRF curves matched. D_n was then multiplied by the square root of the number of data cards contained in the sample to obtain z . For values of z greater than or equal to 0.28 and less than or equal to 1.82, the significance level was obtained from the table contained in Appendix E. For values of z less than 0.28, α was reported as greater than 0.999999. For values of z greater than 1.82, α was reported as less than 0.002645.

E. RESULTS

In an attempt to construct an unbiased set of samples for final testing, an assortment of short stories, novels, poems, reports, and speeches by one hundred and thirty different authors was selected from several anthologies (Appendix F). The samples within this group were equally split between English and Spanish and covered a period from the twelfth century through the present day. Each sample was taken from the beginning of the particular work and was of such length to fill the prescribed text field of ten punch cards.

Each sample was analyzed four times: once utilizing all ten data cards (534 characters); once utilizing the first five data cards (267 characters); once utilizing the first two data cards (107 characters); and once utilizing only the first data card (53 characters).

Furthermore, each of these were processed once under Language-ID-Version-One (Appendix I) which used the space only as a separator and once under Language-ID-Version-Two (Appendix J) which used the space as a character.

In addition to the seven separate tests described in section III. C., three other tests which were evaluated were:

1. a majority vote of all seven tests;
2. a majority vote of the three tests based on Yule's 'K';
3. a majority vote of the four tests based on the Kolmogorov-Smirnov (K-S) test of goodness of fit.

The following tabulation presents the proportion of correct identifications achieved by each of the ten tests under Language-ID-Version-One (V-1) and Language-ID-Version-Two (V-2).

TABLE I
Proportion of Correct Identifications

No. of Cards	K for X_n		K for X_n, X_{n+1}	
	V-1	V-2	V-1	V-2
10	.9538	.6615	.6000	.6231
5	.9000	.6462	.5769	.6384
2	.7615	.5615	.5846	.6385
1	.7077	.5308	.5154	.5923

	K for $X_n, -, X_{n+2}$		K-S for Vowel-Cons.	
	V-1	V-2	V-1	V-2
10	.5308	.5846	1.0000	SAME AS V-1
5	.5077	.5923	1.0000	
2	.4615	.5461	.9923	
1	.5000	.5231	.8770	

	K-S for X_n		K-S for X_n, X_{n+1}	
	V-1	V-2	V-1	V-2
10	1.0000	1.0000	.9923	1.0000
5	.9769	.9769	.9230	.9846
2	.9348	.9348	.8539	.9615
1	.8770	.8693	.7230	.8539

No. of Cards	K-S for $X_n, -, X_{n+2}$		Majority Vote	
	V-1	V-2	V-1	V-2
10	.9307	1.0000	.9846	1.0000
5	.8770	.9769	.9615	1.0000
2	.8154	.9308	.9230	.9615
1	.7521	.8153	.8923	.8923

	Majority of K		Majority of K-S	
	V-1	V-2	V-1	V-2
10	.6385	.6154	.9923	1.0000
5	.6615	.6118	.9461	.9769
2	.6308	.5770	.8923	.9461
1	.5846	.5384	.7846	.8385

IV. CONCLUSIONS

The following observations have been made based upon the results presented in section III. E.

1. Yule's characteristic 'K' by itself was not a reliable measure for identifying the language based on a small sample. Two factors could have affected this measure: first was the close proximity of the mean values when viewed in light of the standard deviation; secondly was that with only twenty-six letters being considered there was too great a chance of repetitiveness which in fact was what Yule was measuring.

2. It was observed that a higher proportion of correct identifications occurred when the sample was analyzed using the space as a character. Accordingly it was concluded that the space rate or the interword structure in each language does contain a measure of intelligence and was useful in identification.

3. The use of the Kolmogorov-Smirnov Goodness of Fit Test when applied to vowel-consonant relationships proved to be the most reliable single test with a significance level of 0.0077 for a sample of approximately 107 characters.

4. No majority vote test was as reliable as the vowel-consonant test but the test based on a majority of the seven individual tests did prove to be the second most reliable test. In view of the fact that this test incorporated the results of three low-reliability Yule's 'K' tests it is subject to some doubt as to how reliable it would prove if the program were expanded to include more than two languages.

Although the experiments using only two languages could not produce conclusive evidence, this research tends to support the belief that each language does possess some statistical characteristic which would allow identification using a short sample of text. Expansion of this research to include additional languages would be useful in determining the reliability of this type of identification system when used with short samples of text.

APPENDIX A. ENGLISH POPULATION VALUES (V-1)

1. Yule's 'K' based on the independent occurrence of each character

X_n :

Mean value 645.079

Standard Deviation 81.479

2. Yule's 'K' based on the joint occurrence of X_n, X_{n+1} :

Mean value 87.168

Standard Deviation 40.207

3. Yule's 'K' based on the joint occurrence of $X_n, -, X_{n+2}$:

Mean value 62.083

Standard Deviation 42.821

TABLE II. ENGLISH JOINT PROBABILITY TABLE FOR VOWEL-CONSONANT

		First Letter	
		Vowel	Consonant
Second Letter	Vowel	.0512	.3568
	Consonant	.3549	.2371

TABLE III. ENGLISH (V-1) INDEPENDENT PROBABILITY OF x_n

x_n	Probability
A	.0800
E	.1214
I	.0710
O	.0774
U	.0284
B	.0138
C	.0313
D	.0418
F	.0222
G	.0217
H	.0502
J	.0017
K	.0072
L	.0445
M	.0255
N	.0731
P	.0210
Q	.0008
R	.0631
S	.0640
T	.0880
V	.0112
W	.0188
X	.0020
Y	.0191
Z	.0008

TABLE IV. ENGLISH (V-1) JOINT PROBABILITY TABLE FOR X_n, X_{n+1}

X_{n+1}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0000	.0072	.0016	.0012	.0010	.0022	.0065	.0022	.0015	.0019	.0096	.0001	.0001
E	.0000	.0050	.0030	.0004	.0013	.0046	.0046	.0058	.0019	.0040	.0271	.0005	.0029
I	.0033	.0020	.0000	.0010	.0010	.0012	.0022	.0038	.0040	.0014	.0086	.0000	.0012
O	.0000	.0007	.0062	.0025	.0001	.0021	.0085	.0038	.0058	.0016	.0065	.0006	.0001
U	.0011	.0002	.0000	.0125	.0000	.0023	.0010	.0013	.0010	.0007	.0019	.0009	.0000
B	.0019	.0002	.0011	.0008	.0006	.0001	.0000	.0001	.0000	.0000	.0001	.0000	.0001
C	.0031	.0040	.0066	.0017	.0017	.0000	.0003	.0001	.0000	.0000	.0000	.0000	.0000
D	.0039	.0133	.0035	.0017	.0007	.0000	.0000	.0005	.0000	.0000	.0000	.0000	.0000
F	.0012	.0012	.0019	.0075	.0002	.0000	.0000	.0000	.0017	.0000	.0000	.0000	.0000
G	.0019	.0009	.0038	.0007	.0016	.0000	.0000	.0000	.0000	.0003	.0000	.0000	.0000
H	.0001	.0002	.0000	.0004	.0000	.0000	.0054	.0001	.0000	.0035	.0000	.0000	.0000
J	.0001	.0000	.0000	.0001	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
K	.0012	.0003	.0004	.0010	.0000	.0000	.0026	.0000	.0000	.0000	.0000	.0000	.0000
L	.0106	.0054	.0051	.0052	.0032	.0020	.0013	.0005	.0007	.0006	.0001	.0000	.0000
M	.0031	.0032	.0037	.0052	.0012	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000
N	.0201	.0150	.0238	.0170	.0050	.0000	.0000	.0002	.0000	.0008	.0002	.0000	.0006
P	.0013	.0016	.0011	.0024	.0017	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Q	.0000	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
R	.0120	.0188	.0033	.0126	.0057	.0007	.0022	.0021	.0019	.0020	.0011	.0000	.0000
S	.0086	.0124	.0085	.0026	.0052	.0002	.0002	.0014	.0001	.0007	.0002	.0000	.0006
T	.0135	.0045	.0093	.0039	.0047	.0001	.0030	.0001	.0018	.0002	.0024	.0000	.0000
V	.0022	.0032	.0035	.0019	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000
W	.0009	.0014	.0000	.0038	.0000	.0000	.0000	.0001	.0001	.0000	.0000	.0000	.0000
X	.0002	.0014	.0006	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0029	.0014	.0000	.0002	.0001	.0015	.0004	.0007	.0001	.0005	.0003	.0000	.0001
Z	.0002	.0001	.0006	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000



TABLE IV - Continued

X_{n+1}	X_n													
	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	
	A	.0053	.0057	.0034	.0030	.0000	.0062	.0035	.0039	.0007	.0047	.0001	.0001	.0002
	E	.0095	.0081	.0090	.0048	.0000	.0174	.0080	.0128	.0102	.0047	.0001	.0012	.0005
	I	.0070	.0042	.0040	.0012	.0000	.0067	.0048	.0109	.0026	.0034	.0002	.0003	.0001
	O	.0035	.0037	.0037	.0035	.0000	.0077	.0033	.0110	.0007	.0031	.0000	.0022	.0001
	U	.0009	.0010	.0006	.0009	.0010	.0011	.0027	.0019	.0000	.0000	.0000	.0000	.0000
	B	.0000	.0008	.0000	.0000	.0000	.0002	.0002	.0001	.0000	.0000	.0000	.0000	.0000
	C	.0000	.0001	.0030	.0000	.0000	.0013	.0013	.0003	.0000	.0000	.0001	.0001	.0000
	D	.0031	.0000	.0151	.0000	.0000	.0021	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	F	.0006	.0000	.0005	.0000	.0000	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000
	G	.0000	.0000	.0114	.0000	.0000	.0019	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	H	.0000	.0000	.0002	.0007	.0000	.0001	.0032	.0300	.0000	.0035	.0001	.0000	.0000
	J	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	K	.0002	.0000	.0008	.0000	.0000	.0011	.0005	.0000	.0000	.0000	.0000	.0000	.0000
	L	.0089	.0001	.0004	.0023	.0000	.0008	.0007	.0011	.0000	.0001	.0000	.0006	.0000

TABLE V. ENGLISH (V-1) JOINT PROBABILITY TABLE FOR X_n , $-$, X_{n+2}

X_{n+2}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0021	.0039	.0057	.0034	.0032	.0008	.0045	.0012	.0008	.0011	.0011	.0000	.0001
E	.0175	.0150	.0168	.0151	.0045	.0025	.0042	.0034	.0030	.0023	.0015	.0001	.0005
I	.0093	.0070	.0068	.0046	.0031	.0015	.0035	.0011	.0016	.0011	.0014	.0001	.0001
O	.0031	.0034	.0014	.0012	.0009	.0007	.0020	.0009	.0025	.0013	.0014	.0001	.0007
U	.0011	.0018	.0025	.0012	.0004	.0012	.0032	.0003	.0011	.0001	.0031	.0000	.0000
B	.0003	.0007	.0002	.0004	.0006	.0002	.0004	.0001	.0001	.0002	.0002	.0002	.0000
C	.0023	.0028	.0014	.0011	.0010	.0014	.0000	.0021	.0015	.0002	.0012	.0002	.0000
D	.0168	.0037	.0025	.0038	.0058	.0004	.0007	.0017	.0006	.0005	.0021	.0001	.0011
F	.0006	.0013	.0006	.0015	.0005	.0004	.0002	.0003	.0006	.0001	.0001	.0000	.0000
G	.0023	.0027	.0125	.0037	.0009	.0007	.0005	.0002	.0001	.0000	.0008	.0000	.0000
H	.0021	.0012	.0070	.0019	.0026	.0000	.0000	.0001	.0000	.0001	.0000	.0002	.0000
J	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
K	.0024	.0009	.0018	.0028	.0003	.0000	.0000	.0000	.0000	.0000	.0002	.0000	.0000
L	.0067	.0031	.0046	.0086	.0027	.0013	.0038	.0025	.0013	.0005	.0034	.0000	.0002
M	.0009	.0011	.0009	.0021	.0001	.0000	.0034	.0008	.0004	.0002	.0030	.0001	.0001
N	.0026	.0041	.0098	.0051	.0012	.0015	.0085	.0032	.0011	.0018	.0067	.0002	.0015
P	.0017	.0031	.0015	.0023	.0011	.0000	.0007	.0005	.0000	.0001	.0012	.0000	.0001
Q	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
R	.0021	.0055	.0005	.0053	.0008	.0016	.0024	.0023	.0078	.0020	.0062	.0004	.0004
S	.0068	.0094	.0049	.0088	.0019	.0008	.0027	.0017	.0008	.0019	.0074	.0009	.0003
T	.0084	.0156	.0070	.0087	.0049	.0031	.0013	.0012	.0004	.0045	.0051	.0000	.0003
V	.0004	.0015	.0010	.0006	.0001	.0002	.0002	.0004	.0016	.0008	.0018	.0000	.0000
W	.0001	.0003	.0002	.0001	.0001	.0001	.0001	.0006	.0003	.0000	.0012	.0001	.0001
X	.0000	.0000	.0000	.0000	.0000	.0002	.0000	.0000	.0001	.0000	.0000	.0000	.0000
Y	.0027	.0018	.0020	.0007	.0002	.0007	.0001	.0015	.0005	.0002	.0013	.0000	.0002
Z	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0001	.0000	.0000

TABLE V - Continued

X_{n+2}	X_n																											
	L	M	N	P	Q	R	S	T	V	W	X	Y	Z															
	A	.0042	.0016	.0023	.0017	.0005	.0047	.0053	.0069	.0001	.0011	.0001	.0014	.0000														
	E	.0038	.0019	.0080	.0045	.0004	.0087	.0061	.0317	.0003	.0019	.0006	.0010	.0001														
	I	.0023	.0012	.0051	.0028	.0003	.0042	.0051	.0054	.0001	.0018	.0002	.0002	.0000														
	O	.0023	.0006	.0028	.0053	.0000	.0022	.0045	.0104	.0001	.0013	.0001	.0001	.0000														
	U	.0007	.0005	.0011	.0004	.0000	.0022	.0024	.0011	.0000	.0011	.0000	.0023	.0000														
	B	.0004	.0001	.0002	.0005	.0000	.0015	.0005	.0005	.0000	.0000	.0000	.0000	.0000														
	C	.0020	.0011	.0007	.0011	.0000	.0045	.0015	.0025	.0006	.0001	.0001	.0000	.0000														
	D	.0021	.0018	.0013	.0009	.0000	.0049	.0021	.0035	.0010	.0004	.0001	.0003	.0002														
	F	.0008	.0001	.0001	.0000	.0000	.0011	.0005	.0002	.0000	.0002	.0000	.0000	.0000														
	G	.0014	.0007	.0005	.0001	.0000	.0017	.0004	.0006	.0000	.0001	.0000	.0000	.0000														
	H	.0004	.0002	.0010	.0000	.0000	.0018	.0006	.0004	.0000	.0000	.0001	.0003	.0000														
	J	.0000	.0002	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000														
	K	.0005	.0011	.0000	.0000	.0000	.0001	.0000	.0005	.0000	.0000	.0000	.0000	.0000														
	L	.0000	.0028	.0030	.0022	.0000	.0030	.0029	.0034	.0013	.0020	.0001	.0001	.0000														
	M	.0011	.0002	.0009	.0000	.0000	.0041	.0025	.0026	.0002	.0002	.0001	.0000	.0000														
	N	.0037	.0070	.0032	.0022	.0000	.0040	.0041	.0052	.0028	.0023	.0001	.0008	.0004														
	P	.0005	.0000	.0002	.0004	.0000	.0020	.0008	.0007	.0000	.0001	.0000	.0000	.0000														
	Q	.0000	.0000	.0000	.0000	.0000	.0002	.0000	.0001	.0000	.0000	.0000	.0000	.0000														
	R	.0039	.0043	.0057	.0055	.0000	.0005	.0040	.0094	.0047	.0037	.0002	.0004	.0002														
	S	.0042	.0035	.0042	.0016	.0000	.0059	.0015	.0018	.0015	.0031	.0002	.0002	.0000														
	T	.0038	.0028	.0063	.0013	.0000	.0036	.0011	.0016	.0007	.0028	.0000	.0003	.0001														
	V	.0008	.0004	.0007	.0000	.0000	.0017	.0014	.0009	.0001	.0004	.0000	.0000	.0000														
	W	.0013	.0000	.0023	.0002	.0000	.0011	.0002	.0003	.0000	.0000	.0000	.0000	.0000														
	X	.0000	.0001	.0003	.0000	.0000	.0000	.0000	.0003	.0000	.0000	.0000	.0000	.0000														
Y	.0018	.0006	.0022	.0009	.0000	.0025	.0010	.0017	.0000	.0011	.0000	.0000	.0000															
Z	.0002	.0001	.0004	.0000	.0000	.0001	.0001	.0001	.0000	.0000	.0000	.0000	.0000															

APPENDIX B. SPANISH POPULATION VALUES (V-1)

1. Yule's 'K' based on the independent occurrence of each character

X_n :

Mean value 751.447

Standard Deviation 82.328

2. Yule's 'K' based on the joint occurrence of X_n, X_{n+1} :

Mean value 104.553

Standard Deviation 38.094

3. Yule's 'K' based on the joint occurrence of $X_n, -, X_{n+2}$:

Mean value 72.680

Standard Deviation 35.210

TABLE VI. SPANISH JOINT PROBABILITY TABLE FOR VOWEL-CONSONANT

		First Letter	
		Vowel	Consonant
Second Letter	Vowel	.0616	.4469
	Consonant	.3713	.1202

TABLE VII. SPANISH (V-1) INDEPENDENT PROBABILITY OF X_n

X_n	Probability
A	.1269
E	.1311
I	.0690
O	.0929
U	.0409
B	.0137
C	.0448
D	.0470
F	.0072
G	.0111
H	.0087
J	.0041
K	.0006
L	.0568
M	.0285
N	.0747
P	.0247
Q	.0087
R	.0619
S	.0752
T	.0464
V	.0114
W	.0000
X	.0014
Y	.0089
Z	.0037

TABLE VIII. SPANISH (V-1) JOINT PROBABILITY TABLE FOR X_n, X_{n+1}

X_{n+1}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0000	.0017	.0115	.0001	.0032	.0055	.0106	.0076	.0009	.0027	.0050	.0012	.0001
E	.0003	.0003	.0108	.0000	.0146	.0013	.0049	.0278	.0015	.0015	.0012	.0013	.0001
I	.0009	.0009	.0000	.0004	.0029	.0033	.0132	.0066	.0026	.0016	.0013	.0001	.0002
O	.0003	.0011	.0104	.0003	.0002	.0013	.0133	.0119	.0015	.0024	.0028	.0019	.0001
U	.0012	.0006	.0002	.0000	.0000	.0008	.0047	.0017	.0014	.0028	.0006	.0006	.0000
B	.0066	.0006	.0017	.0020	.0008	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
C	.0081	.0062	.0070	.0033	.0022	.0000	.0011	.0000	.0000	.0000	.0000	.0000	.0000
D	.0101	.0037	.0053	.0026	.0015	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
F	.0005	.0013	.0008	.0007	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
G	.0008	.0029	.0021	.0011	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
H	.0003	.0000	.0000	.0000	.0000	.0000	.0035	.0000	.0000	.0000	.0000	.0000	.0000
J	.0009	.0014	.0004	.0004	.0003	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
K	.0000	.0001	.0000	.0001	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000
L	.0116	.0139	.0048	.0037	.0020	.0018	.0009	.0000	.0004	.0003	.0000	.0000	.0001
M	.0054	.0035	.0041	.0053	.0012	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
N	.0176	.0302	.0095	.0164	.0111	.0000	.0001	.0000	.0000	.0002	.0001	.0000	.0000
P	.0015	.0011	.0005	.0014	.0008	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Q	.0006	.0003	.0002	.0001	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
R	.0173	.0162	.0028	.0102	.0031	.0026	.0009	.0011	.0008	.0023	.0000	.0000	.0000
S	.0157	.0234	.0049	.0230	.0032	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0001
T	.0038	.0037	.0035	.0021	.0011	.0001	.0029	.0000	.0000	.0000	.0000	.0000	.0000
V	.0009	.0015	.0024	.0023	.0003	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0001	.0014	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0009	.0002	.0000	.0004	.0006	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Z	.0009	.0014	.0008	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000



x_n x_{n+1}

TABLE IX. SPANISH (V-1) JOINT PROBABILITY TABLE FOR X_n , $-$, X_{n+2}

X_{n+2}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0196	.0140	.0169	.0095	.0106	.0031	.0094	.0024	.0007	.0023	.0000	.0001	.0000
E	.0128	.0113	.0056	.0084	.0030	.0047	.0052	.0022	.0020	.0020	.0003	.0001	.0000
I	.0171	.0134	.0097	.0091	.0036	.0016	.0043	.0005	.0003	.0015	.0000	.0000	.0000
O	.0158	.0108	.0103	.0084	.0053	.0009	.0086	.0015	.0003	.0016	.0000	.0000	.0001
U	.0034	.0034	.0026	.0016	.0005	.0002	.0015	.0001	.0001	.0003	.0000	.0000	.0000
B	.0011	.0005	.0003	.0012	.0003	.0003	.0010	.0007	.0002	.0001	.0025	.0002	.0000
C	.0024	.0054	.0020	.0012	.0020	.0004	.0008	.0028	.0023	.0004	.0019	.0002	.0000
D	.0035	.0027	.0016	.0021	.0022	.0003	.0025	.0035	.0000	.0005	.0001	.0002	.0000
F	.0002	.0006	.0006	.0004	.0001	.0001	.0001	.0007	.0000	.0000	.0000	.0002	.0000
G	.0021	.0005	.0007	.0011	.0009	.0000	.0002	.0001	.0002	.0000	.0001	.0001	.0000
H	.0008	.0008	.0001	.0008	.0011	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
J	.0000	.0001	.0003	.0003	.0001	.0005	.0000	.0005	.0000	.0001	.0002	.0000	.0000
K	.0002	.0001	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
L	.0041	.0025	.0045	.0008	.0023	.0008	.0037	.0042	.0005	.0006	.0006	.0000	.0002
M	.0021	.0018	.0028	.0010	.0003	.0003	.0057	.0012	.0002	.0002	.0014	.0005	.0000
N	.0014	.0020	.0160	.0003	.0031	.0017	.0112	.0035	.0017	.0027	.0011	.0007	.0000
P	.0004	.0051	.0014	.0019	.0005	.0000	.0013	.0002	.0000	.0000	.0001	.0000	.0000
Q	.0007	.0001	.0002	.0003	.0003	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000
R	.0041	.0040	.0037	.0032	.0031	.0021	.0055	.0043	.0018	.0022	.0013	.0008	.0000
S	.0021	.0028	.0047	.0021	.0028	.0008	.0050	.0086	.0006	.0009	.0016	.0010	.0000
T	.0100	.0248	.0055	.0041	.0036	.0009	.0007	.0012	.0005	.0005	.0001	.0003	.0000
V	.0005	.0013	.0007	.0006	.0008	.0000	.0002	.0008	.0000	.0000	.0001	.0007	.0000
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0000	.0001	.0000	.0000	.0000	.0001	.0001	.0000	.0000	.0000	.0005	.0000	.0001
Z	.0009	.0006	.0009	.0000	.0001	.0001	.0002	.0000	.0000	.0001	.0003	.0000	.0000

TABLE IX - Continued

X_n															
	L	M	N	P	Q	R	S	T	V	W	X	Y	Z		
A	.0046	.0015	.0090	.0021	.0000	.0093	.0078	.0052	.0007	.0000	.0000	.0000	.0001		
E	.0029	.0024	.0149	.0057	.0121	.0076	.0048	.0036	.0016	.0000	.0002	.0000	.0000		
I	.0016	.0015	.0073	.0023	.0024	.0047	.0026	.0010	.0003	.0000	.0001	.0000	.0000		
O	.0029	.0015	.0113	.0032	.0000	.0077	.0050	.0034	.0007	.0000	.0001	.0001	.0003		
U	.0009	.0001	.0029	.0003	.0000	.0010	.0024	.0007	.0000	.0000	.0000	.0000	.0000		
B	.0017	.0002	.0003	.0004	.0000	.0023	.0014	.0016	.0002	.0000	.0000	.0000	.0002		
C	.0032	.0016	.0036	.0024	.0000	.0047	.0010	.0042	.0013	.0000	.0001	.0000	.0001		
D	.0016	.0023	.0024	.0017	.0000	.0025	.0013	.0049	.0013	.0000	.0000	.0001	.0003		
F	.0005	.0000	.0002	.0000	.0000	.0008	.0002	.0006	.0000	.0000	.0000	.0000	.0000		
G	.0019	.0005	.0003	.0001	.0000	.0019	.0015	.0010	.0002	.0000	.0001	.0000	.0000		
H	.0000	.0000	.0003	.0000	.0000	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000		
J	.0003	.0006	.0001	.0001	.0000	.0004	.0004	.0001	.0001	.0000	.0000	.0000	.0000		
K	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000		
L	.0002	.0044	.0020	.0018	.0000	.0013	.0035	.0030	.0011	.0000	.0005	.0000	.0001		
M	.0023	.0003	.0012	.0001	.0000	.0028	.0013	.0029	.0004	.0000	.0003	.0000	.0000		
N	.0030	.0107	.0020	.0022	.0000	.0066	.0050	.0065	.0019	.0000	.0001	.0002	.0006		
P	.0002	.0000	.0001	.0001	.0000	.0022	.0009	.0002	.0000	.0000	.0000	.0000	.0000		
Q	.0002	.0001	.0000	.0001	.0000	.0001	.0001	.0000	.0002	.0000	.0000	.0000	.0000		
R	.0027	.0047	.0058	.0128	.0000	.0015	.0051	.0072	.0027	.0000	.0004	.0005	.0004		
S	.0154	.0053	.0079	.0026	.0000	.0102	.0040	.0072	.0022	.0000	.0003	.0002	.0007		
T	.0016	.0021	.0015	.0007	.0000	.0024	.0011	.0010	.0003	.0000	.0000	.0000	.0000		
V	.0010	.0004	.0010	.0000	.0000	.0008	.0013	.0016	.0003	.0000	.0000	.0000	.0000		
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000		
X	.0000	.0001	.0000	.0000	.0000	.0001	.0001	.0001	.0000	.0000	.0000	.0000	.0000		
Y	.0003	.0009	.0000	.0001	.0000	.0003	.0002	.0002	.0001	.0000	.0000	.0000	.0000		
Z	.0006	.0000	.0002	.0002	.0000	.0008	.0000	.0002	.0006	.0000	.0000	.0000	.0000		

APPENDIX C. ENGLISH POPULATION VALUES (V-2)

1. Yule's 'K' based on the independent occurrence of each character

X_n :

Mean value	747.770
------------	---------

Standard Deviation	95.973
--------------------	--------

2. Yule's 'K' based on the joint occurrence of X_n , X_{n+1} :

Mean value	87.930
------------	--------

Standard Deviation	31.618
--------------------	--------

3. Yule's 'K' based on the joint occurrence of X_n , -, X_{n+2} :

Mean value	73.139
------------	--------

Standard Deviation	27.466
--------------------	--------

4.. Vowel-Consonant relationships are displayed in Table II.

TABLE X. ENGLISH (V-2) INDEPENDENT PROBABILITY OF x_n

x_n	Probability
A	.0659
E	.1000
I	.0585
O	.0638
U	.0234
B	.0114
C	.0258
D	.0344
F	.0183
G	.0179
H	.0414
J	.0014
K	.0059
L	.0367
M	.0210
N	.0603
P	.0173
Q	.0006
R	.0520
S	.0528
T	.0725
V	.0093
W	.0155
X	.0016
Y	.0157
Z	.0007
Space	.1760

TABLE XI. ENGLISH (V-2) JOINT PROBABILITY TABLE FOR X_n, X_{n+1}

X_n		A	E	I	O	U	B	C	D	F	G	H	J	K
X_{n+1}	A	.0000	.0047	.0010	.0008	.0007	.0014	.0042	.0014	.0010	.0012	.0062	.0001	.0001
	E	.0000	.0033	.0020	.0002	.0008	.0030	.0030	.0037	.0012	.0026	.0176	.0003	.0019
	I	.0022	.0013	.0000	.0007	.0006	.0008	.0014	.0025	.0025	.0009	.0056	.0000	.0008
	O	.0000	.0005	.0040	.0016	.0000	.0014	.0055	.0025	.0037	.0010	.0042	.0004	.0001
	U	.0007	.0001	.0000	.0081	.0000	.0015	.0017	.0008	.0007	.0005	.0012	.0006	.0000
	B	.0013	.0002	.0007	.0005	.0004	.0001	.0000	.0000	.0000	.0000	.0001	.0000	.0000
	C	.0020	.0026	.0043	.0011	.0011	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000
	D	.0025	.0086	.0023	.0011	.0005	.0000	.0000	.0003	.0000	.0000	.0000	.0000	.0000
	F	.0008	.0008	.0013	.0049	.0001	.0000	.0000	.0000	.0011	.0000	.0000	.0000	.0000
	G	.0012	.0006	.0025	.0005	.0010	.0000	.0000	.0001	.0000	.0002	.0000	.0000	.0000
	H	.0000	.0001	.0000	.0002	.0000	.0000	.0035	.0000	.0000	.0023	.0000	.0000	.0000
	J	.0001	.0000	.0000	.0001	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	K	.0008	.0002	.0003	.0006	.0000	.0000	.0017	.0000	.0000	.0000	.0000	.0000	.0000
	L	.0069	.0035	.0033	.0034	.0020	.0013	.0009	.0003	.0005	.0004	.0001	.0000	.0002
	M	.0020	.0020	.0024	.0034	.0008	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000
	N	.0130	.0097	.0154	.0110	.0032	.0000	.0000	.0001	.0000	.0005	.0002	.0000	.0004
	P	.0008	.0011	.0007	.0016	.0011	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	Q	.0000	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	R	.0077	.0122	.0022	.0082	.0037	.0005	.0015	.0014	.0013	.0013	.0007	.0000	.0000
	S	.0056	.0080	.0055	.0017	.0034	.0001	.0001	.0009	.0001	.0005	.0001	.0000	.0004
	T	.0087	.0029	.0060	.0025	.0030	.0001	.0020	.0000	.0012	.0001	.0015	.0000	.0000
	V	.0014	.0021	.0023	.0012	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
	W	.0006	.0009	.0000	.0024	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
	X	.0002	.0009	.0004	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	Y	.0019	.0009	.0000	.0001	.0001	.0010	.0003	.0004	.0000	.0003	.0002	.0000	.0000
	Z	.0001	.0000	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	Space	.0055	.0329	.0016	.0077	.0008	.0002	.0009	.0194	.0051	.0062	.0036	.0000	.0021

TABLE XI - Continued

		X_n														
		L	M	N	P	Q	R	S	T	V	W	X	Y	Z	Space	
X_{n+1}	A	.0034	.0037	.0022	.0019	.0000	.0040	.0023	.0025	.0005	.0030	.0001	.0001	.0001	.0195	
	E	.0062	.0053	.0058	.0031	.0000	.0113	.0052	.0083	.0066	.0031	.0001	.0008	.0003	.0045	
	I	.0045	.0027	.0026	.0008	.0000	.0043	.0031	.0071	.0017	.0022	.0001	.0002	.0001	.0100	
	O	.0023	.0024	.0024	.0022	.0000	.0050	.0021	.0072	.0004	.0020	.0000	.0014	.0000	.0114	
	U	.0006	.0006	.0004	.0006	.0006	.0007	.0017	.0012	.0000	.0000	.0000	.0000	.0000	.0020	
	B	.0000	.0005	.0000	.0000	.0000	.0001	.0002	.0000	.0000	.0000	.0000	.0001	.0000	.0072	
	C	.0000	.0001	.0019	.0000	.0000	.0009	.0008	.0002	.0000	.0000	.0001	.0000	.0000	.0105	
	D	.0020	.0000	.0098	.0000	.0000	.0013	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0059	
	F	.0004	.0000	.0003	.0000	.0000	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0084	
	G	.0000	.0000	.0074	.0000	.0000	.0012	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0032	
	H	.0000	.0000	.0001	.0004	.0000	.0001	.0021	.0194	.0000	.0023	.0001	.0000	.0000	.0107	
	J	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0012	
	K	.0002	.0000	.0005	.0000	.0000	.0007	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0007	
	L	.0057	.0001	.0003	.0015	.0000	.0005	.0005	.0007	.0000	.0001	.0000	.0004	.0000	.0043	
	M	.0001	.0005	.0003	.0000	.0000	.0008	.0007	.0002	.0000	.0000	.0000	.0000	.0000	.0077	
	N	.0000	.0000	.0005	.0000	.0000	.0012	.0001	.0000	.0000	.0005	.0000	.0001	.0000	.0043	
	P	.0003	.0016	.0001	.0010	.0000	.0003	.0011	.0000	.0000	.0000	.0003	.0000	.0000	.0074	
	Q	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0003	
	R	.0001	.0001	.0000	.0030	.0000	.0007	.0000	.0026	.0000	.0001	.0000	.0002	.0000	.0046	
	S	.0010	.0005	.0025	.0004	.0000	.0039	.0023	.0019	.0000	.0003	.0000	.0004	.0000	.0132	
	T	.0005	.0000	.0065	.0004	.0000	.0023	.0066	.0011	.0000	.0000	.0005	.0001	.0000	.0264	
	V	.0002	.0000	.0006	.0000	.0000	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0011	
	W	.0000	.0000	.0001	.0000	.0000	.0001	.0003	.0011	.0000	.0000	.0000	.0000	.0000	.0098	
	X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
	Y	.0027	.0005	.0007	.0002	.0000	.0014	.0002	.0027	.0001	.0000	.0000	.0000	.0000	.0019	
	Z	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	
	Space	.0064	.0026	.0153	.0016	.0000	.0108	.0230	.0161	.0000	.0019	.0004	.0118	.0000	.0000	

TABLE XII. ENGLISH (V-2) JOINT PROBABILITY TABLE FOR X_n , -, X_{n+2}

X_n		A	E	I	O	U	B	C	D	F	G	H	J	K
X_{n+2}	A	.0011	.0048	.0029	.0022	.0016	.0004	.0023	.0029	.0009	.0013	.0009	.0000	.0004
	E	.0084	.0079	.0081	.0075	.0021	.0012	.0020	.0021	.0015	.0012	.0008	.0000	.0003
	I	.0045	.0054	.0033	.0026	.0015	.0007	.0018	.0017	.0010	.0010	.0009	.0000	.0002
	O	.0015	.0037	.0007	.0009	.0005	.0004	.0011	.0015	.0015	.0011	.0009	.0000	.0005
	U	.0005	.0014	.0012	.0007	.0002	.0006	.0015	.0004	.0006	.0001	.0015	.0000	.0001
	B	.0004	.0016	.0002	.0006	.0003	.0001	.0002	.0013	.0001	.0003	.0002	.0001	.0001
	C	.0017	.0036	.0008	.0010	.0006	.0007	.0001	.0018	.0011	.0004	.0007	.0001	.0001
	D	.0085	.0028	.0013	.0022	.0028	.0002	.0004	.0016	.0004	.0004	.0011	.0001	.0006
	F	.0007	.0020	.0003	.0009	.0002	.0002	.0002	.0011	.0005	.0003	.0002	.0000	.0002
	G	.0012	.0021	.0060	.0021	.0005	.0003	.0002	.0003	.0002	.0001	.0004	.0000	.0000
	H	.0012	.0028	.0035	.0014	.0013	.0000	.0001	.0014	.0003	.0004	.0003	.0001	.0001
	J	.0000	.0002	.0000	.0001	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000
	K	.0012	.0006	.0009	.0014	.0002	.0000	.0000	.0001	.0000	.0000	.0001	.0000	.0000
	L	.0035	.0024	.0023	.0043	.0013	.0007	.0018	.0017	.0017	.0004	.0017	.0000	.0002
	M	.0009	.0022	.0006	.0015	.0001	.0000	.0016	.0011	.0004	.0003	.0017	.0001	.0001
	N	.0015	.0029	.0047	.0026	.0006	.0007	.0041	.0021	.0006	.0010	.0033	.0001	.0007
	P	.0013	.0032	.0007	.0015	.0005	.0000	.0004	.0008	.0003	.0003	.0007	.0000	.0001
	Q	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
	R	.0012	.0038	.0003	.0029	.0004	.0008	.0012	.0016	.0038	.0011	.0031	.0002	.0003
	S	.0039	.0072	.0024	.0051	.0010	.0004	.0014	.0021	.0007	.0012	.0038	.0004	.0003
	T	.0043	.0112	.0035	.0051	.0024	.0015	.0007	.0039	.0016	.0034	.0031	.0000	.0004
	V	.0002	.0009	.0005	.0003	.0001	.0001	.0001	.0004	.0008	.0004	.0009	.0000	.0000
	W	.0004	.0024	.0003	.0004	.0001	.0001	.0001	.0011	.0003	.0003	.0009	.0000	.0002
	X	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0001	.0000	.0000	.0000	.0000
	Y	.0014	.0012	.0010	.0005	.0001	.0003	.0001	.0009	.0003	.0002	.0007	.0000	.0002
	Z	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	Space	.0161	.0239	.0129	.0164	.0049	.0020	.0045	.0024	.0007	.0026	.0136	.0001	.0011

TABLE XII - Continued

X_n															
	L	M	N	P	Q	R	S	T	V	W	X	Y	Z	Space	
A	.0027	.0011	.0031	.0010	.0003	.0034	.0064	.0050	.0001	.0006	.0001	.0020	.0000	.0186	
E	.0021	.0010	.0045	.0022	.0002	.0045	.0034	.0156	.0001	.0010	.0003	.0008	.0000	.0211	
I	.0013	.0007	.0031	.0015	.0002	.0025	.0041	.0036	.0001	.0010	.0001	.0009	.0000	.0151	
O	.0016	.0005	.0023	.0027	.0000	.0018	.0039	.0065	.0001	.0007	.0001	.0008	.0000	.0288	
U	.0004	.0003	.0007	.0002	.0000	.0011	.0013	.0008	.0000	.0006	.0000	.0013	.0000	.0081	
B	.0005	.0001	.0006	.0003	.0000	.0011	.0012	.0008	.0000	.0001	.0000	.0005	.0000	.0008	
C	.0014	.0007	.0012	.0006	.0000	.0029	.0019	.0021	.0003	.0002	.0001	.0007	.0000	.0012	
D	.0012	.0010	.0012	.0005	.0000	.0027	.0016	.0022	.0005	.0003	.0001	.0006	.0001	.0005	
F	.0008	.0002	.0007	.0001	.0000	.0010	.0015	.0008	.0000	.0002	.0001	.0009	.0000	.0056	
G	.0009	.0004	.0003	.0001	.0000	.0010	.0005	.0006	.0000	.0001	.0000	.0002	.0000	.0003	
H	.0005	.0002	.0015	.0002	.0000	.0015	.0014	.0013	.0000	.0002	.0001	.0007	.0000	.0211	
J	.0001	.0001	.0001	.0000	.0000	.0002	.0002	.0001	.0000	.0000	.0000	.0001	.0000	.0000	
K	.0003	.0005	.0001	.0000	.0000	.0001	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0001	
L	.0002	.0014	.0017	.0011	.0000	.0017	.0018	.0021	.0007	.0010	.0001	.0003	.0000	.0038	
M	.0007	.0002	.0010	.0001	.0000	.0025	.0019	.0018	.0001	.0002	.0001	.0005	.0000	.0015	
N	.0019	.0034	.0019	.0011	.0000	.0022	.0025	.0028	.0013	.0012	.0001	.0008	.0002	.0162	
P	.0006	.0001	.0007	.0002	.0000	.0015	.0012	.0008	.0000	.0001	.0000	.0004	.0000	.0018	
Q	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0002	
R	.0021	.0021	.0030	.0027	.0000	.0005	.0024	.0048	.0023	.0018	.0001	.0005	.0001	.0090	
S	.0025	.0019	.0029	.0009	.0000	.0036	.0021	.0022	.0007	.0016	.0001	.0012	.0000	.0031	
T	.0025	.0019	.0065	.0008	.0000	.0035	.0038	.0034	.0004	.0016	.0001	.0017	.0001	.0054	
V	.0004	.0002	.0004	.0000	.0000	.0009	.0008	.0005	.0001	.0002	.0000	.0001	.0000	.0011	
W	.0010	.0002	.0016	.0002	.0000	.0011	.0017	.0010	.0000	.0001	.0000	.0006	.0000	.0016	
X	.0000	.0000	.0002	.0000	.0000	.0000	.0004	.0002	.0000	.0000	.0000	.0000	.0000	.0007	
Y	.0010	.0003	.0013	.0004	.0000	.0013	.0006	.0011	.0000	.0007	.0002	.0001	.0000	.0021	
Z	.0010	.0000	.0002	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
Space	.0100	.0027	.0197	.0006	.0000	.0094	.0061	.0123	.0027	.0024	.0001	.0003	.0001	.0084	
X_{n+2}															

APPENDIX D. SPANISH POPULATION VALUES (V-2)

1. Yule's 'K' based on the independent occurrence of each character

X_n :

Mean value	806.481
------------	---------

Standard Deviation	77.810
--------------------	--------

2. Yule's 'K' based on the joint occurrence of X_n, X_{n+1} :

Mean value	108.784
------------	---------

Standard Deviation	30.665
--------------------	--------

3. Yule's 'K' based on the joint occurrence of $X_n, -, X_{n+2}$:

Mean value	83.553
------------	--------

Standard Deviation	22.886
--------------------	--------

4. Vowel-Consonant relationships are displayed in Table VI.

TABLE XIII. SPANISH (V-2) INDEPENDENT PROBABILITY OF x_n

x_n	Probability
A	.1053
E	.1088
I	.0573
O	.0771
U	.0339
B	.0114
C	.0372
D	.0390
F	.0060
G	.0092
H	.0072
J	.0034
K	.0005
L	.0471
M	.0236
N	.0620
P	.0205
Q	.0073
R	.0514
S	.0625
T	.0385
V	.0095
W	.0000
X	.0011
Y	.0074
Z	.0031
Space	.1699

TABLE XIV. SPANISH (V-2) JOINT PROBABILITY TABLE FOR X_n, X_{n+1}

X_{n+1}	X_n												
	A	E	I	O	U	B	C	D	F	G	H	J	K
A	.0000	.0011	.0076	.0001	.0021	.0036	.0070	.0050	.0006	.0018	.0033	.0008	.0000
E	.0002	.0002	.0071	.0000	.0096	.0008	.0033	.0183	.0010	.0010	.0008	.0009	.0001
I	.0006	.0006	.0000	.0002	.0019	.0022	.0087	.0044	.0017	.0011	.0009	.0001	.0002
O	.0002	.0007	.0068	.0002	.0002	.0008	.0088	.0079	.0010	.0016	.0018	.0012	.0000
U	.0008	.0004	.0001	.0000	.0000	.0005	.0031	.0011	.0009	.0019	.0004	.0004	.0000
B	.0043	.0004	.0011	.0013	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
C	.0054	.0041	.0046	.0022	.0014	.0000	.0007	.0000	.0000	.0000	.0000	.0000	.0000
D	.0067	.0025	.0035	.0017	.0010	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
F	.0003	.0009	.0005	.0005	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
G	.0005	.0019	.0014	.0007	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
H	.0002	.0000	.0000	.0000	.0000	.0000	.0023	.0000	.0000	.0000	.0000	.0000	.0000
J	.0006	.0009	.0003	.0003	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000
K	.0000	.0001	.0000	.0001	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000
L	.0077	.0092	.0032	.0024	.0013	.0012	.0006	.0000	.0003	.0002	.0000	.0000	.0001
M	.0036	.0023	.0027	.0035	.0008	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
N	.0116	.0200	.0063	.0108	.0073	.0000	.0001	.0000	.0000	.0002	.0000	.0000	.0000
P	.0010	.0007	.0003	.0009	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Q	.0004	.0002	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
R	.0114	.0107	.0018	.0067	.0020	.0017	.0006	.0007	.0005	.0015	.0000	.0000	.0000
S	.0104	.0155	.0033	.0152	.0021	.0003	.0000	.0000	.0000	.0000	.0000	.0000	.0001
T	.0025	.0025	.0023	.0014	.0007	.0001	.0019	.0000	.0000	.0000	.0000	.0000	.0000
V	.0006	.0010	.0016	.0015	.0002	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000
W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
X	.0000	.0010	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Y	.0006	.0002	.0000	.0003	.0004	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Z	.0006	.0010	.0005	.0001	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
Space	.0352	.0311	.0022	.0268	.0013	.0000	.0000	.0014	.0000	.0000	.0000	.0000	.0001

TABLE XIV - Continued

X_n		L	M	N	P	Q	R	S	T	V	W	X	Y	Z	Space
X_{n+1}	A	.0133	.0050	.0072	.0052	.0000	.0103	.0034	.0093	.0019	.0000	.0001	.0006	.0014	.0146
	E	.0050	.0057	.0031	.0034	.0000	.0085	.0073	.0096	.0032	.0000	.0000	.0003	.0000	.0196
	I	.0039	.0041	.0030	.0013	.0000	.0059	.0045	.0054	.0032	.0000	.0003	.0000	.0000	.0032
	O	.0075	.0035	.0062	.0041	.0000	.0072	.0042	.0075	.0010	.0000	.0000	.0006	.0007	.0034
	U	.0007	.0015	.0016	.0014	.0072	.0011	.0030	.0016	.0001	.0000	.0000	.0001	.0001	.0058
	B	.0001	.0012	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0022
	C	.0002	.0000	.0032	.0000	.0000	.0008	.0012	.0000	.0000	.0000	.0001	.0000	.0001	.0134
	D	.0002	.0000	.0031	.0000	.0000	.0010	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0193
	F	.0001	.0000	.0006	.0000	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0028
	G	.0005	.0000	.0008	.0000	.0000	.0009	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0021
	H	.0000	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0045
	J	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0010
	K	.0000	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	L	.0028	.0000	.0000	.0015	.0000	.0006	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0159
	M	.0004	.0000	.0001	.0000	.0000	.0018	.0006	.0000	.0000	.0000	.0000	.0000	.0000	.0078
	N	.0000	.0000	.0003	.0000	.0000	.0004	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0049
	P	.0001	.0025	.0000	.0000	.0000	.0002	.0017	.0000	.0000	.0000	.0003	.0000	.0000	.0123
	Q	.0000	.0000	.0004	.0000	.0000	.0002	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0057
	R	.0001	.0000	.0000	.0032	.0000	.0014	.0000	.0050	.0000	.0000	.0000	.0000	.0000	.0039
	S	.0003	.0000	.0017	.0000	.0000	.0013	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0124
	T	.0008	.0000	.0116	.0001	.0000	.0020	.0065	.0000	.0000	.0000	.0002	.0000	.0000	.0061
	V	.0004	.0000	.0006	.0000	.0000	.0005	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0031
	W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	Y	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0058
	Z	.0000	.0000	.0005	.0000	.0000	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0001
	Space	.0106	.0000	.0178	.0001	.0000	.0071	.0295	.0001	.0001	.0000	.0000	.0058	.0007	n/c

TABLE XV. SPANISH (V-2) JOINT PROBABILITY TABLE FOR X_n , X_{n+1} , X_{n+2}

x_n														
	A	E	I	O	U	B	C	D	F	G	H	J	K	
x_{n+2}	A	.0123	.0095	.0086	.0074	.0054	.0016	.0047	.0013	.0004	.0012	.0000	.0001	.0000
	E	.0102	.0090	.0031	.0078	.0016	.0024	.0026	.0013	.0010	.0010	.0001	.0001	.0000
	I	.0092	.0074	.0049	.0051	.0019	.0008	.0022	.0003	.0002	.0007	.0000	.0000	.0000
	O	.0088	.0059	.0052	.0045	.0027	.0004	.0043	.0008	.0001	.0008	.0000	.0000	.0001
	U	.0032	.0029	.0014	.0018	.0003	.0001	.0007	.0000	.0001	.0001	.0000	.0000	.0000
	B	.0011	.0007	.0002	.0008	.0001	.0001	.0005	.0004	.0001	.0000	.0013	.0001	.0000
	C	.0042	.0051	.0011	.0021	.0011	.0002	.0004	.0015	.0012	.0002	.0010	.0001	.0000
	D	.0056	.0037	.0009	.0051	.0012	.0002	.0012	.0020	.0000	.0003	.0001	.0001	.0000
	F	.0008	.0007	.0003	.0005	.0001	.0000	.0001	.0004	.0000	.0000	.0000	.0001	.0000
	G	.0016	.0007	.0004	.0008	.0005	.0000	.0001	.0000	.0001	.0000	.0001	.0000	.0000
	H	.0011	.0014	.0001	.0014	.0006	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	J	.0002	.0002	.0002	.0002	.0000	.0003	.0000	.0003	.0000	.0000	.0001	.0000	.0000
	K	.0001	.0001	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	L	.0049	.0059	.0025	.0022	.0012	.0004	.0019	.0021	.0002	.0003	.0003	.0000	.0001
	M	.0029	.0024	.0016	.0018	.0003	.0001	.0028	.0007	.0001	.0001	.0007	.0002	.0000
	N	.0015	.0020	.0082	.0010	.0016	.0008	.0056	.0018	.0008	.0013	.0005	.0003	.0000
	P	.0030	.0045	.0009	.0025	.0004	.0000	.0007	.0002	.0000	.0000	.0000	.0000	.0000
	Q	.0015	.0010	.0002	.0015	.0002	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	R	.0029	.0028	.0019	.0020	.0016	.0010	.0027	.0022	.0009	.0011	.0006	.0004	.0000
	S	.0033	.0039	.0026	.0029	.0015	.0004	.0025	.0044	.0003	.0005	.0008	.0005	.0000
	T	.0061	.0133	.0028	.0029	.0019	.0004	.0004	.0006	.0002	.0003	.0000	.0001	.0000
	V	.0012	.0013	.0004	.0007	.0005	.0000	.0001	.0004	.0000	.0000	.0000	.0004	.0000
	W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	X	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000
	Y	.0015	.0008	.0001	.0010	.0000	.0000	.0001	.0001	.0000	.0000	.0002	.0000	.0001
	Z	.0004	.0003	.0004	.0000	.0001	.0000	.0001	.0000	.0000	.0000	.0002	.0000	.0000
	Space	.0176	.0235	.0095	.0210	.0094	.0019	.0036	.0181	.0002	.0013	.0011	.0008	.0001

TABLE XV - Continued

		X_n														
		L	M	N	P	Q	R	S	T	V	W	X	Y	Z	Space	
X_{n+2}	A	.0032	.0008	.0061	.0011	.0000	.0055	.0064	.0026	.0004	.0000	.0000	.0005	.0001	.0263	
	E	.0023	.0012	.0100	.0028	.0060	.0048	.0054	.0018	.0008	.0000	.0001	.0006	.0001	.0326	
	I	.0010	.0008	.0040	.0012	.0012	.0024	.0018	.0005	.0002	.0000	.0000	.0000	.0000	.0116	
	O	.0017	.0007	.0060	.0016	.0000	.0039	.0032	.0017	.0004	.0000	.0000	.0002	.0002	.0239	
	U	.0005	.0001	.0024	.0002	.0000	.0009	.0018	.0003	.0000	.0000	.0000	.0002	.0000	.0171	
	B	.0011	.0001	.0004	.0002	.0000	.0012	.0009	.0008	.0001	.0000	.0000	.0001	.0001	.0009	
	C	.0029	.0008	.0032	.0012	.0000	.0028	.0029	.0021	.0007	.0000	.0000	.0005	.0001	.0018	
	D	.0020	.0012	.0026	.0009	.0000	.0019	.0054	.0025	.0006	.0000	.0000	.0006	.0002	.0010	
	F	.0045	.0000	.0004	.0000	.0000	.0006	.0006	.0003	.0000	.0000	.0000	.0001	.0000	.0006	
	G	.0011	.0002	.0004	.0000	.0000	.0010	.0010	.0005	.0001	.0000	.0000	.0001	.0000	.0004	
	H	.0003	.0000	.0005	.0000	.0000	.0002	.0008	.0000	.0000	.0000	.0000	.0002	.0000	.0007	
	J	.0003	.0003	.0002	.0000	.0000	.0003	.0004	.0001	.0000	.0000	.0000	.0000	.0000	.0003	
	K	.0001	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
	L	.0005	.0022	.0032	.0009	.0000	.0017	.0035	.0015	.0006	.0000	.0002	.0007	.0001	.0100	
	M	.0016	.0002	.0014	.0000	.0000	.0016	.0016	.0015	.0002	.0000	.0002	.0004	.0001	.0013	
	N	.0020	.0053	.0014	.0011	.0000	.0034	.0033	.0033	.0010	.0000	.0001	.0004	.0003	.0152	
	P	.0014	.0000	.0010	.0000	.0000	.0014	.0029	.0001	.0000	.0000	.0000	.0003	.0001	.0008	
	Q	.0003	.0001	.0005	.0000	.0000	.0004	.0010	.0000	.0001	.0000	.0000	.0001	.0001	.0004	
	R	.0018	.0023	.0033	.0064	.0000	.0009	.0032	.0036	.0013	.0000	.0002	.0004	.0002	.0076	
	S	.0083	.0027	.0052	.0013	.0000	.0056	.0042	.0036	.0011	.0000	.0001	.0008	.0004	.0056	
	T	.0014	.0010	.0016	.0004	.0000	.0015	.0016	.0005	.0001	.0000	.0000	.0002	.0000	.0011	
	V	.0007	.0002	.0008	.0000	.0000	.0005	.0011	.0008	.0001	.0000	.0000	.0000	.0000	.0003	
	W	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	.0000	
	X	.0000	.0000	.0000	.0000	.0000	.0001	.0000	.0001	.0000	.0000	.0000	.0000	.0000	.0009	
	Y	.0003	.0005	.0004	.0001	.0000	.0004	.0015	.0001	.0000	.0000	.0000	.0001	.0002	.0001	
	Z	.0003	.0000	.0001	.0001	.0000	.0004	.0000	.0001	.0003	.0000	.0000	.0000	.0000	.0001	
	Space	.0117	.0031	.0071	.0010	.0000	.0082	.0076	.0103	.0014	.0000	.0000	.0010	.0010	.0095	

APPENDIX E. TABLE XVI - Significance Levels for Kolmogorov-Smirnov Test

Z	α	Z	α	Z	α	Z	α	Z	α
0.28	.999999	0.59	.877240	0.90	.392730	1.21	.106970	1.52	.019690
0.29	.999996	0.60	.864282	0.91	.379072	1.22	.101896	1.53	.018524
0.30	.999991	0.61	.850771	0.92	.365714	1.23	.097028	1.54	.017422
0.31	.999979	0.62	.836775	0.93	.352662	1.24	.092352	1.55	.016378
0.32	.999954	0.63	.822247	0.94	.339918	1.25	.087868	1.56	.015390
0.33	.999909	0.64	.807333	0.95	.327484	1.26	.083568	1.57	.014456
0.34	.999829	0.65	.792013	0.96	.315364	1.27	.079444	1.58	.013574
0.35	.999697	0.66	.776363	0.97	.303556	1.28	.075495	1.59	.012740
0.36	.999489	0.67	.760418	0.98	.292060	1.29	.071712	1.60	.011952
0.37	.999174	0.68	.744220	0.99	.280874	1.30	.068092	1.61	.011209
0.38	.998715	0.69	.727811	1.00	.270000	1.31	.064630	1.62	.010508
0.39	.998071	0.70	.711235	1.01	.259434	1.32	.061318	1.63	.009846
0.40	.997192	0.71	.694529	1.02	.249174	1.33	.058152	1.64	.009223
0.41	.996028	0.72	.677735	1.03	.239220	1.34	.055128	1.65	.008636
0.42	.994524	0.73	.660887	1.04	.229566	1.35	.052244	1.66	.008083
0.43	.992623	0.74	.644019	1.05	.220206	1.36	.049488	1.67	.007562
0.44	.990270	0.75	.627167	1.06	.211140	1.37	.046858	1.68	.007072
0.45	.987410	0.76	.610360	1.07	.202364	1.38	.044350	1.69	.006611
0.46	.983995	0.77	.593628	1.08	.193872	1.39	.041960	1.70	.006177
0.47	.979978	0.78	.576998	1.09	.185658	1.40	.039682	1.71	.005770
0.48	.975318	0.79	.560495	1.10	.177718	1.41	.037514	1.72	.005388
0.49	.969983	0.80	.544143	1.11	.170050	1.42	.035448	1.73	.005028
0.50	.963645	0.81	.527959	1.12	.162644	1.43	.033484	1.74	.004691
0.51	.957186	0.82	.511970	1.13	.155498	1.44	.031618	1.75	.004375
0.52	.949694	0.83	.496192	1.14	.148606	1.45	.029842	1.76	.004078
0.53	.941466	0.84	.480634	1.15	.141962	1.46	.028154	1.77	.003800
0.54	.932503	0.85	.465318	1.16	.135558	1.47	.026552	1.78	.003540
0.55	.922817	0.86	.450256	1.17	.129388	1.48	.025030	1.79	.003296
0.56	.912423	0.87	.435454	1.18	.123452	1.49	.023588	1.80	.003068
0.57	.901344	0.88	.420930	1.19	.117742	1.50	.022218	1.81	.002845
0.58	.889605	0.89	.406684	1.20	.112250	1.51	.020920	1.82	.002645

APPENDIX F. FINAL TEST SOURCES

From Introduction to Literature: Stories [Ref. 10]:

Anderson, Sherwood, Death in the Woods.
Benét, Stephen Vincent, By The Waters of Babylon.
Crane, Stephen, The Open Boat.
Curley, Daniel, A Story of Love, Etc.
Edmonds, Walter D., Death of Red Peril.
Fitzgerald, F. Scott, Absolution.
Forster, E. M., The Other Side of the Hedge.
Galsworthy, John, Quality.
Hardy, Thomas, The Three Strangers.
Porter, Katherine Anne, The Grave.
Schwartz, Delmore, In Dreams Begin Responsibilities.
Steele, Wilbur Daniel, Footfalls.
Steinbeck, John, from The Red Pony: The Gift.
Stevenson, Robert Louis, Markheim.
White, E. B., The Dove.

From Five Centuries of Spanish Literature [Ref. 11]:

Anonymous, Lazarillo de Tormes.
Cervantes Saavedra, Miguel De, El ingenioso hidalgo, don Quijote de la Mancha.
Manuel, Don Juan, El Conde Lucanor.

From Antología Venezolana [Ref. 12]:

Alfonzo, Alfredo Armas, La hora y punto.
Arráiz, Antonio, El mar es como un potro vigoroso.

Coll, Pedro Emilio, El diente roto.

Croce, Arturo, Angustia apagada.

Diaz Rodríguez, Manuel, Mi Alma era una mina abandonada.

Diaz Sanchez, Ramon, Poderes de la mas alta valia.

Diaz Solis, Gustavo, La efigie.

Fombona, Rufino Blanco, El catire.

Gallegos, Romulo, Canaima.

Garmendia, Julio, Eladia.

Guaramato, Oscar, Vecindad.

La Parra, Teresa De, Vicente Cochocho.

Marquez Salas, Antonio, Como Dios!.

Meneses, Guillermo, El compañero Juan Ruiz.

Mujica, Hector, Las tres ventanas.

Núñez, Enrique Bernardo, Vocchi.

Palacios, Lucila, Cap. XX (Cubil).

Picon Salas, Mariano, Josefita.

Pocaterra, Jose Rafael, Patria, la mestiza.

Pardo, Isaac J., El domador de potros.

Pardon, Julian, II (Primavera nocturna).

Rivas Mijares, Humberto, El murado.

Rosales, Julio, El mejor rabula.

Ruiz, Jose Fabbiani, A orillas del viejo rio.

Silva, Miguel Otero, Fiebre.

Stolk, Gloria, Cap. I (Amargo fondo).

Urbaneja Achelpohl, Luis M., Ovejon...!.

Uslar Pietri, Arturo, X (Las lanzas coloradas).

From An Anthology of Spanish American Literature [Ref. 13]:

Altamirano, Ignacio Manuel, El dia de muertos.

Bastamente Carlos Inga, Calixto (Concolorcorvo), El lazarillo de ciegos caminantes.

Bello, Andres, Silva a la Argicultura de la Zona Torrida.

Blanco-Fombona, Rufino, El conquistador español del siglo XVI.

Bolivar, Simon, Carta a un caballero que tomaba gran interes en la causa republicana en la america del sur.

Chocano, Jose Santos, La Epopeya Del Pacifico.

Cortes, Hernan, The Death of Moctezuma and the Tragic Retreat of the Spaniards from the Aztec Capital the Night of June 30, 1520.

Dario, Ruben, Azul (El rey burgues).

Diaz.del Castillo, Bernal, The Imprisonment of Moctezuma.

Fernandez de Lizardi, Jose Joaquin, El periquillo sarniento.

Garcia Calderon, Francisco, La creacion de un continente.

Garcilaso de la Vega, El Inca, Comentarios reales de los incas.

Gonzalez Prada, Manuel, Discurso en el politeama.

Gutierrez Najera, Manuel, La Novela Del Tranvia.

Hernandez, Jose, Martin Fierro.

Juana Ines de la Cruz, Sor, Respuesta a Sor Filotea de la Cruz.

Latorre, Mariano, Chilenos del mar (El piloto Oyarzo).

Lillo, Baldomero, El chiflon del diablo.

Lopez Albuja, Enrique, Cuentos andinos (Como habla la coca).

Marti, Jose, Nuestra America.

Mitre, Bartolome, La historia de San Martin.

Montalvo, Juan, Washington y Bolivar.

Nervo, Amado, Una esperanza.

Palma, Ricardo, Las orejas del Alcalde.

Quiroga, Horacio, El desierto.

Reyes, Alfonso, Vision de Anahuac.

Rodo, Jose Enrique, Ariel.

Rojas, Manuel, Hombres del sur (El cachorro).

Rojas, Ricardo, La argentinidad.

Sarmiento, Domingo Faustino, Facundo.

Siguenza y Gongora, Carlos de, Infortunios de Alonso Rameriz.

Ugarte, Manuel, El destino de un continente (La neuva Roma).

Vasconcelos, Jose, La raza cosmica.

Viana, Javier de, Leña seca (El domador).

From The Creative Reader [Ref. 14]:

Algren, Nelson, The Moon of the Arfy Darfy.

Amster, L. J., Center of Gravity.

Auchincloss, Louis, Power in Trust.

Berriault, Gina, The Birthday Party.

Boyle, Kay, The Ballet of Central Park.

Buechler, James, John Sobieski Runs.

Calisher, Hortense, Little Did I Know.

Connell, Evan S., Jr., The Suicide.

Cozzens, James Gould, One Hundred Ladies.

Elkin, Stanley, Perlmutter at the East Pole.

Gary, Romain, A Humanist.

Gold, Hebert, Dance of the Divorced.

Goyen, William, Figure over the Town.

Greene, Graham, The Root of All Evil.

Malamud, Bernard, The Refugee.

McCullers, Carson, Sucker.

Miller, Warren, Chaos, Disorder and the Late Show.

Moravia, Alberto, A Tough Nut.
O'Connor, Frank, A Life of Your Own.
O'Faolain, Sean, One Man, One Boat, One Girl.
Pynchon, Thomas, The Secret Integration.
Saroyan, William, Boys and Girls Together.
Seager, Allan, No More Roses.
Shaw, Irwin, The Inhabitants of Venus.
Singer, Isaac Bashevis, Esther Kreindel The Second.
Swados, Harvey, A Story For Teddy.
Updike, John, The Lucid Eye in Silver Town.
White, Robin, Walker's Peak.
Williams, Thomas, The Old Dancers.
Wood, Malcolm, The Appraiser.

From Best Modern Short Stories [Ref. 15]:

Aiken, Conrad, Mr. Arcularis.
Audry, Colette, The Gloves.
Bierce, Ambrose, An Occurrence at Owl Creek Bridge.
Bowen, Elizabeth, The Demon Lover.
Conrad, Hoseph, Amy Foster.
Faulkner, William, A Rose For Emily.
Granberry, Edwin, A Trip to Czardis.
Hawthorne, Nathaniel, My Kinsman, Major Molineux.
Hemingway, Ernest, A Clean, Well-Lighted Place.
James, Henry, Paste.
Joyce, James, A Little Cloud.
Lawrence, D. H., The Horse Dealer's Daughter.
Lowry, Malcolm, Strange Comfort Afforded by the Profession.

Mansfield, Katherine, The Fly.

Maupassant, Guy de, Two Friends.

Poe, Edgar Allan, Ligeia.

Taylor, Elizabeth, The First Death of Her Life.

Thurber, James, The Catbird Seat.

Welty, A Still Moment.

West, Jessamyn, Sixteen.

002250	77	VC	TOTAL	PIC	9(8)	VALUE	ZEROS.	ANAL	-ONE
002300	77	AB	PIC	99,	COMP,	SYNC.		ANAL	-ONE
002350	77	BN	PIC	99,	COMP,	SYNC.		ANAL	-ONE
002400	77	NE	PIC	99,	CONS	PIC	9(7),	ANAL	-ONE
002450	77	F-	CONS	-VOW			COMP,	ANAL	-ONE
002500	77	F-	CONS	-VOW			SYNC,	ANAL	-ONE
002550	77	F-	CONS	-VOW			VALUE	ANAL	-ONE
002600	77	F-	CONS	-VOW			ZEROS.	ANAL	-ONE
002650	77	FO	L-FREQ	-TOTAL	PIC	9(8),	COMP,	ANAL	-ONE
002700	77	S	LF	-TOTAL	PIC	9(8),	SYNC.	ANAL	-ONE
002750	77	IC	RF	-TOTAL	PIC	9(8),	COMP,	ANAL	-ONE
002800	77	CR	F	-TOTAL	PIC	9(5),	CCMP-3.	ANAL	-ONE
002850	77	CA	RD	-NR	PIC	9(6),	COMP	ANAL	-ONE
002900	77	LM	X	PIC	99,	COMP,	SYNC.	ANAL	-ONE
002950	77	XY	PIC	99,	COMP,	SYNC.		ANAL	-ONE
003000	77	Y	PIC	99,	COMP,	SYNC.		ANAL	-ONE
003050	77	W	S	-TABLE	-NR,	PIC	99,	ANAL	-ONE
003100	77	LT	R	-TABLE	-NR,	PIC	99,	ANAL	-ONE
003150	01	L5	5	5	5	5	5	ANAL	-ONE
003200		5	5	5	5	5	5	ANAL	-ONE
003250		5	5	5	5	5	5	ANAL	-ONE
003300		5	5	5	5	5	5	ANAL	-ONE
003350		5	5	5	5	5	5	ANAL	-ONE
003400		5	5	5	5	5	5	ANAL	-ONE
003450		5	5	5	5	5	5	ANAL	-ONE
003500		5	5	5	5	5	5	ANAL	-ONE
003550		5	5	5	5	5	5	ANAL	-ONE
003600		5	5	5	5	5	5	ANAL	-ONE
003650		5	5	5	5	5	5	ANAL	-ONE
003700		5	5	5	5	5	5	ANAL	-ONE
003750		5	5	5	5	5	5	ANAL	-ONE
003800		5	5	5	5	5	5	ANAL	-ONE
003850		5	5	5	5	5	5	ANAL	-ONE
003900		5	5	5	5	5	5	ANAL	-ONE
003950		5	5	5	5	5	5	ANAL	-ONE
004000		5	5	5	5	5	5	ANAL	-ONE
004050		5	5	5	5	5	5	ANAL	-ONE
004100		5	5	5	5	5	5	ANAL	-ONE
004150		5	5	5	5	5	5	ANAL	-ONE
004200		5	5	5	5	5	5	ANAL	-ONE
004250		5	5	5	5	5	5	ANAL	-ONE
004300		5	5	5	5	5	5	ANAL	-ONE
004350		5	5	5	5	5	5	ANAL	-ONE
004400		5	5	5	5	5	5	ANAL	-ONE
004450		5	5	5	5	5	5	ANAL	-ONE
004500		5	5	5	5	5	5	ANAL	-ONE
004550		5	5	5	5	5	5	ANAL	-ONE
004600		5	5	5	5	5	5	ANAL	-ONE

004650	05	FILLER	PIC	X	VALUE 'Y'.	ANAL-ONE
004700	05	FILLER	PIC	X	VALUE 'Z'.	ANAL-ONE
004750	05	FILLER	PIC	X	VALUE '6'.	ANAL-ONE
004800	05	FILLER	PIC	X	VALUE '7'.	ANAL-ONE
004850	01	LTR-TABLE	REDEFINES	LTR-TABLE-VALUES.		ANAL-ONE
004900	05	LTR	PIC	X,	OCCURS 33 TIMES.	ANAL-ONE
004950	01	IND-FREQ-TABLE	PIC	9(6),	COMP, SYNC, OCCURS 33 TIMES.	ANAL-ONE
005000	05	ILF	PIC	9V9(5),	COMP, SYNC, OCCURS 33 TIMES.	ANAL-ONE
005050	01	IND-PROB-TABLE	PIC	9V9(5),	COMP, SYNC, OCCURS 33 TIMES.	ANAL-ONE
005100	05	ILP	PIC	9V9(5),	COMP, SYNC, OCCURS 33 TIMES.	ANAL-ONE
005150	01	DEPENDENCY-FREQ-TABLE	PIC	9(6),	COMP, SYNC, OCCURS 66 TIMES.	ANAL-ONE
005200	05	FST-LTR	OCCURS	PIC 9(6),	COMP, SYNC, OCCURS 66 TIMES.	ANAL-ONE
005250	01	DEPENDENCY-FREQ-TABLE	PIC	9(6),	COMP, SYNC, OCCURS 66 TIMES.	ANAL-ONE
005300	05	IO FOL	PROB	PIC 9V9(5),	COMP, SYNC, OCCURS 66 TIMES.	ANAL-ONE
005350	05	GIVEN-LTR	OCCURS	33 TIMES.		ANAL-ONE
005400	01	IO FOL-PROB	PIC	9V9(5),	COMP, SYNC, OCCURS 66 TIMES.	ANAL-ONE
005450	05	OUTPT-ONE-A	PIC	X(43) VALUE SPACES.		ANAL-ONE
005500	05	FILLER	PIC	X(48) VALUE 'FREQUENCY/INDEPENDENT PROBABILITY'		ANAL-ONE
005550	05	FILLER	PIC	X(48) VALUE 'FREQUENCY/INDEPENDENT PROBABILITY'		ANAL-ONE
005600	-	CF EACH LETTER				ANAL-ONE
005650	05	FILLER	PIC	X(42) VALUE SPACES.		ANAL-ONE
005700	01	LANG-LINE	PIC	X(54) VALUE SPACES.		ANAL-ONE
005750	05	FILLER	PIC	X(11) VALUE 'LANGUAGE:.'		ANAL-ONE
005800	05	FILLER	PIC	X(11) VALUE 'LANGUAGE:.'		ANAL-ONE
005850	05	LANG-DESIG	PIC	X(30) VALUE SPACES.		ANAL-ONE
005900	05	FILLER	PIC	X(38) VALUE SPACES.		ANAL-ONE
005950	01	OUTPT-ONE-C	PIC	X(37) VALUE SPACES.		ANAL-ONE
006000	05	FILLER	PIC	X(25) VALUE 'LETTER.'		ANAL-ONE
006050	05	FILLER	PIC	X(25) VALUE 'FREQUENCY.'		ANAL-ONE
006100	05	FILLER	PIC	X(25) VALUE 'PROBABILITY.'		ANAL-ONE
006150	05	FILLER	PIC	X(46) VALUE 'PROBABILITY.'		ANAL-ONE
006200	01	OUTPT-ONE-B	PIC	X(43) VALUE SPACES.		ANAL-ONE
006250	05	FILLER	PIC	X(48) VALUE ALL.		ANAL-ONE
006300	05	FILLER	PIC	X(42) VALUE SPACES.		ANAL-ONE
006350	05	FILLER	PIC	X(42) VALUE SPACES.		ANAL-ONE
006400	01	LANG-UNDERLINE	PIC	X(65) VALUE SPACES.		ANAL-ONE
006450	05	FILLER	PIC	X(30) VALUE ALL.		ANAL-ONE
006500	05	FILLER	PIC	X(38) VALUE SPACES.		ANAL-ONE
006550	05	FILLER	PIC	X(38) VALUE SPACES.		ANAL-ONE
006600	01	OUTPT-ONE-D	PIC	X(37) VALUE SPACES.		ANAL-ONE
006650	05	FILLER	PIC	X(61) VALUE ALL.		ANAL-ONE
006700	05	FILLER	PIC	X(35) VALUE SPACES.		ANAL-ONE
006750	05	FILLER	PIC	X(35) VALUE SPACES.		ANAL-ONE
006800	01	OUTPT-ONE-E	PIC	X(40) VALUE SPACES.		ANAL-ONE
006850	05	FILLER	PIC	X(40) VALUE SPACES.		ANAL-ONE
006900	05	LTR-ONE	PIC	X(22) VALUE SPACES.		ANAL-ONE
006950	05	FILLER	PIC	X(22) VALUE SPACES.		ANAL-ONE
007000	05	FREQ-ONE	PIC	ZZZ,ZZ9.		ANAL-ONE

007050	05	FILLER	PIC	X(19)	VALUE	SPACES.	ANAL-ONE
007100	05	PROB-ONE	PIC	9.99999.			ANAL-ONE
007150	05	FILLER	PIC	X(37)	VALUE	SPACES.	ANAL-ONE
007200	01	OUTPT	-ONE-F				ANAL-ONE
007250	05	FILLER	PIC	X(38)	VALUE	SPACES.	ANAL-ONE
007300	05	FILLER	PIC	X(22)	VALUE	TOTAL.	ANAL-ONE
007350	05	ONE-F	ILF-TLT	PIC	ZZ,ZZZ,ZZ9.		ANAL-ONE
007400	05	FILLER	PIC	X(63)	VALUE	SPACES.	ANAL-ONE
007450	01	OUTPT	-TWO-A				ANAL-ONE
007500	05	FILLER	PIC	X(46)	VALUE	SPACES.	ANAL-ONE
007550	05	FILLER	PIC	X(87)	VALUE	FREQUENCY AND PROBABILITY OF (X(2	ANAL-ONE
007600	-01)/X(1))					ANAL-ONE
007650	01	OUTPT	-TWO-THREE-B				ANAL-ONE
007700	05	FILLER	PIC	X(46)	VALUE	SPACES.	ANAL-ONE
007750	05	FILLER	PIC	X(40)	VALUE	ALL.	ANAL-ONE
007800	05	FILLER	PIC	X(47)	VALUE	SPACES.	ANAL-ONE
007850	01	FIRST	-LTR-OUTPT				ANAL-ONE
007900	05	FILLER	PIC	X(31)	VALUE	SPACES.	ANAL-ONE
007950	05	FILLER	PIC	X(102)	VALUE	*****FIRST LETTER, X(1)*****	ANAL-ONE
008000	-01	*****					ANAL-ONE
008050	01	FIRST	-LTR-TITLES				ANAL-ONE
008100	05	FILLER	PIC	X(14)	VALUE	SPACES.	ANAL-ONE
008150	05	FIRST	-LTR	PIC	X(9)	OCCURS 13 TIMES.	ANAL-ONE
008200	05	FILLER	PIC	XX	VALUE	SPACES.	ANAL-ONE
008250	01	DEPN	-FREQ-LINE				ANAL-ONE
008300	05	FILLER	PIC	X(5)	VALUE	SPACES.	ANAL-ONE
008350	05	FREQ	-ROW-LTR	PIC	X.		ANAL-ONE
008400	05	FILLER	PIC	XXX	VALUE	(F).	ANAL-ONE
008450	05	DEPN	-FREQ	PIC	BBZZZ,ZZ9	OCCURS 13 TIMES.	ANAL-ONE
008500	05	FILLER	PIC	X(7)	VALUE	SPACES.	ANAL-ONE
008550	01	DEPN	-PROB-LINE				ANAL-ONE
008600	05	FILLER	PIC	X(5)	VALUE	SPACES.	ANAL-ONE
008650	05	PROB	-ROW-LTR	PIC	X.		ANAL-ONE
008700	05	FILLER	PIC	XX	VALUE	(P).	ANAL-ONE
008750	05	DEPN	-PROB	PIC	BB9.99999	OCCURS 13 TIMES.	ANAL-ONE
008800	05	FILLER	PIC	X(7)	VALUE	SPACES.	ANAL-ONE
008850	01	OUTPT	-THREE-A				ANAL-ONE
008900	05	FILLER	PIC	X(46)	VALUE	SPACES.	ANAL-ONE
008950	05	FILLER	PIC	X(87)	VALUE	FREQUENCY AND PROBABILITY OF (X(3	ANAL-ONE
009000	-01)/X(1))					ANAL-ONE
009050	01	OUTPT	-FOUR-A				ANAL-ONE
009100	05	FILLER	PIC	X(47)	VALUE	SPACES.	ANAL-ONE
009150	05	FILLER	PIC	X(33)	VALUE	P(VOWEL/VOWEL)).	ANAL-ONE
009200	05	P	-VOW-VOW	PIC	9.99999.		ANAL-ONE
009250	05	FILLER	PIC	X(46)	VALUE	SPACES.	ANAL-ONE
009300	01	OUTPT	-FOUR-B				ANAL-ONE
009350	05	FILLER	PIC	X(47)	VALUE	SPACES.	ANAL-ONE
009400	05	FILLER	PIC	X(33)	VALUE	P(CONSONANT/VOWEL)).	ANAL-ONE

009450	05	P-CONS-VOW	PIC	9.99999.			ANAL-ONE
009500	05	FILLER-PIC	X(46)	VALUE	SPACES.		ANAL-ONE
009550	01	OUTPT-FOUR-C					ANAL-ONE
009600	05	FILLER-PIC	X(47)	VALUE	SPACES.		ANAL-ONE
009650	05	FILLER-PIC	X(33)	VALUE	'P(VOWEL/CONSONANT)'. .		ANAL-ONE
009700	05	P-VOW-CONS	PIC	9.99999.			ANAL-ONE
009750	05	FILLER-PIC	X(46)	VALUE	SPACES.		ANAL-ONE
009800	01	OUTPT-FOUR-D					ANAL-ONE
009850	05	FILLER-PIC	X(47)	VALUE	SPACES.		ANAL-ONE
009900	05	FILLER-PIC	X(33)	VALUE	'P(CONSONANT/CONSONANT)'. .		ANAL-ONE
009950	05	P-CCNS-CUN	PIC	9.99999.			ANAL-ONE
010000	05	FILLER-PIC	X(46)	VALUE	SPACES.		ANAL-ONE
010050	01	OUTPT-FOUR-E					ANAL-ONE
010100	05	FILLER-PIC	X(11)	VALUE	SPACES.		ANAL-ONE
010150	05	FILLER-PIC	X(32)	VALUE	'YULE-S		ANAL-ONE
010200	05	OUTPT-FOUR-LANG	PIC	X(30).	CHARACTERISTIC, K, FOR:'. .		ANAL-ONE
010250	05	YULES-K	PIC	Z(12)9.999.			ANAL-ONE
010300	05	FILLER-PIC	X(40)	VALUE	SPACES.		ANAL-ONE
010350	01	OUTPT-FOUR-F					ANAL-ONE
010400	05	FILLER-PIC	X.				ANAL-ONE
010450	05	FILLER-PIC	X(94)	VALUE	'KOLMOGOROV-SMIRNOV CUMULATED REL		ANAL-ONE
010500	-	ATIVE	FREQUENCY	TABLE	FOR INDEPENDENT		ANAL-ONE
010550	-	IVE			LETTER OCCURRENCES IN:		ANAL-ONE
010600	05	OUTPT-FOUR-KS-LANG	PIC	X(30).			ANAL-ONE
010650	05	FILLER-PIC	X(9)	VALUE	SPACES.		ANAL-ONE
010700	01	OUTPT-FOUR-G					ANAL-ONE
010750	05	FILLER-PIC	X(37)	VALUE	SPACES.		ANAL-ONE
010800	05	FILLER-PIC	X(25)	VALUE	'LETTER'. .		ANAL-ONE
010850	05	FILLER-PIC	X(71)	VALUE	'CUM. REL. FREQ.'. .		ANAL-ONE
010900	01	KS-LINE					ANAL-ONE
010950	05	FILLER-PIC	X(40)	VALUE	SPACES.		ANAL-ONE
011000	05	KS-LTR	PIC	X.			ANAL-ONE
011050	05	FILLER-PIC	X(25)	VALUE	SPACES.		ANAL-ONE
011100	05	KS-CREF	PIC	9.99999.			ANAL-ONE
011150	05	FILLER-PIC	X(60)	VALUE	SPACES.		ANAL-ONE
011200	01	ERROR-LINE					ANAL-ONE
011250	05	FILLER-PIC	X.				ANAL-ONE
011300	05	FILLER-PIC	X(28)	VALUE	'OVERFLOW ERROR AT STATEMENT '. .		ANAL-ONE
011350	05	ER-STEP	PIC	9(6).			ANAL-ONE
011400	05	FILLER-PIC	X(21)	VALUE	' CARD IN PROCESS WAS '. .		ANAL-ONE
011450	05	ER-CARD	PIC	9(6).			ANAL-ONE
011500	05	FILLER-PIC	X(71)	VALUE	ALL '*'. .		ANAL-ONE
011550	01	DATA-FORM					ANAL-ONE
011600	05	CC-PIC	X,	OCCURS	73 TIMES.		ANAL-ONE
011650	05	FILLER-PIC	X(7).				ANAL-ONE
011700	01	OUTPT-FIVE-A					ANAL-ONE
011750	05	FILLER-PIC	X(49)	VALUE	SPACES.		ANAL-ONE
011800	05	FILLER-PIC	X(37)	VALUE	'JOINT PROBABILITY TABLE FOR X(1)-ANAL-ONE		ANAL-ONE


```

01 1850-      'X(2)'.
05 FILLER PIC X(47) VALUE SPACES.
01 1900      OUTPT-FIVE-B.
05 FILLER PIC X(11) VALUE SPACES.
05 FILLER PIC X(25) VALUE 'YULE-S K FOR X(1)-X(2):'.
01 1950      YULES-K-2 PIC Z(12)9.999.
05 FILLER PIC X(80) VALUE SPACES.
05 2000      OUTPT-SIX-A.
05 FILLER PIC X(49) VALUE SPACES.
05 FILLER PIC X(37) VALUE 'JOINT PROBABILITY TABLE FOR X(1)-X(3):'.
01 2050      'X(3)'.
05 FILLER PIC X(47) VALUE SPACES.
05 2100      OUTPT-SIX-B.
05 FILLER PIC X(11) VALUE SPACES.
05 FILLER PIC X(25) VALUE 'YULE-S K FOR X(1)-X(3):'.
01 2150      YULES-K-3 PIC Z(12)9.999.
05 FILLER PIC X(80) VALUE SPACES.
05 2200      CARD-FILL.
05 TEST PIC X.999.
05 SEQ-NR PIC X(10).
05 LANG-ID PIC 9V9(5), OCCURS 11 TIMES.
05 CRF-OUT PIC 9V9(5).
PROCEDURE DIVISION.
START-UP.
OPEN INPUT POPULATION, OUTPUT LANG-ANAL, CARD-FILE.
BEGIN.
READ POPULATION AT END GO TO EOJ.
IF NEW-LANG PERFORM CLOSE-OUT-RT THRU INITIALIZE-RT.
PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.
ADD 1 TO CARD-NR ON SIZE ERROR GO TO E-1.
MOVE 3 TO N.
DATA-COLL.
IF CC (N) = SPACE GO TO BLANK-COL-STEP.
MOVE 1 TO A.
CK-CNE.
IF CC (N) = LTR (A) GO TO PROC-TWO.
IF A = 33 GO TO COL-STEP.
ADD 1 TO A.
GO TO CK-CNE.
PROC-TWO.
ADD 1 TO ILF (A) ON SIZE ERROR GO TO E-2.
COMPUTE M = N + 1.
IF CC (M) = SPACE GO TO COL-STEP.
IF CC (M) = '-' AND M = 73 PERFORM CARD-STEP-ONE.
MOVE 1 TO B.
CK-TWO.
IF CC (M) = LTR (B) GO TO PROC-THREE.
IF B = 33 GO TO COL-STEP.

```


016650	MOVE CC (71) TO CC (1).	ANAL-ONE
016700	MOVE CC (72) TO CC (2).	ANAL-ONE
016750	READ POPULATION AT END EOJ.	ANAL-ONE
016800	PERFORM DAT-SHIFT ON SIZE ERROR GO TO E-10.	ANAL-ONE
016850	ADD 1 TO CARD-NR + 2.	ANAL-ONE
016900	COMPUTE L = N + 2.	ANAL-ONE
016950	CLOSE-OUT-RT.	ANAL-ONE
017000	IF CARD-NR = 0 GO TO INITIALIZE-RT.	ANAL-ONE
017050	MOVE ZEROS TO ILF-TOTAL.	ANAL-ONE
017100	PERFORM IND-TOTAL-CALC VARYING A FROM 1 BY 1 UNTIL A > 33.	ANAL-ONE
017150	PERFORM IND-PROB-CALC VARYING A FROM 1 BY 1 UNTIL A > 33.	ANAL-ONE
017200	PERFORM OUTPT-CNE-PRER.	ANAL-ONE
017250	PERFORM OUTPT-CNE-PRINT VARYING A FROM 1 BY 1 UNTIL A > 33.	ANAL-ONE
017300	MOVE ILF-TOTAL TO ONE-F-ILF-TLT.	ANAL-ONE
017350	WRITE PRINT-LINE-CNE-F BEFORE 1.	ANAL-ONE
017400	PERFORM DEPNT-TWO-CALC VARYING A FROM 1 BY 1 UNTIL A > 33.	ANAL-ONE
017450	PERFORM OUTPT-TWO-PRER.	ANAL-ONE
017500	PERFORM COL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-ONE
017550	PERFORM PRINT-LINE-FROMP-A FIRST-LTR-TITLES BEFORE 2.	ANAL-ONE
017600	PERFORM OUTPT-TWO-DUMP-A VARYING B FROM 1 BY 1 UNTIL B > 23.	ANAL-ONE
017650	PERFORM OUTPT-TWO-PRER.	ANAL-ONE
017700	PERFORM COL-TITLES-B VARYING A FROM 14 BY 1 UNTIL A > 26.	ANAL-ONE
017750	PERFORM PRINT-LINE-FROMP-B FIRST-LTR-TITLES BEFORE 2.	ANAL-ONE
017800	PERFORM OUTPT-TWO-DUMP-B VARYING B FROM 1 BY 1 UNTIL B > 23.	ANAL-ONE
017850	PERFORM OUTPT-TWO-PRER.	ANAL-ONE
017900	PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-ONE
017950	PERFORM COL-TITLES-C VARYING A FROM 27 BY 1 UNTIL A > 33.	ANAL-ONE
018000	PERFORM PRINT-LINE-FROMP-C FIRST-LTR-TITLES BEFORE 2.	ANAL-ONE
018050	PERFORM OUTPT-TWO-DUMP-C VARYING B FROM 1 BY 1 UNTIL B > 23.	ANAL-ONE
018100	PERFORM OUTPT-TWO-PRER.	ANAL-ONE
018150	PERFORM COL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-ONE
018200	PERFORM PRINT-LINE-FROMP-A FIRST-LTR-TITLES BEFORE 2.	ANAL-ONE
018250	PERFORM OUTPT-TWO-DUMP-A VARYING B FROM 24 BY 1 UNTIL B > 33.	ANAL-ONE
018300	PERFORM OUTPT-TWO-PRER.	ANAL-ONE
018350	PERFORM COL-TITLES-B VARYING A FROM 14 BY 1 UNTIL A > 26.	ANAL-ONE
018400	PERFORM PRINT-LINE-FROMP-B FIRST-LTR-TITLES BEFORE 2.	ANAL-ONE
018450	PERFORM OUTPT-TWO-DUMP-B VARYING B FROM 24 BY 1 UNTIL B > 33.	ANAL-ONE
018500	PERFORM OUTPT-TWO-PRER.	ANAL-ONE
018550	PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-ONE
018600	PERFORM COL-TITLES-C VARYING A FROM 27 BY 1 UNTIL A > 33.	ANAL-ONE
018650	PERFORM PRINT-LINE-FROMP-C FIRST-LTR-TITLES BEFORE 2.	ANAL-ONE
018700	PERFORM OUTPT-TWO-DUMP-C VARYING B FROM 24 BY 1 UNTIL B > 33.	ANAL-ONE
018750	PERFORM OUTPT-TWO-PRER.	ANAL-ONE
018800	PERFORM DEPNT-TREE-CALC VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-ONE
018850	PERFORM COL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-ONE
018900	PERFORM PRINT-LINE-FROMP-A FIRST-LTR-TITLES BEFORE 2.	ANAL-ONE
018950	PERFORM OUTPT-TWO-DUMP-A VARYING B FROM 34 BY 1 UNTIL B > 56.	ANAL-ONE
019000	PERFORM OUTPT-TREE-PRER.	ANAL-ONE


```

019050 PERFORM COL-TITLES-B VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-ONE
019100 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
019150 PERFORM OUTPT-TWO-DUMP-B VARYING B FROM 34 BY 1 UNTIL B > 56. ANAL-ONE
019200 PERFORM OUTPT-THREE-PRP. ANAL-ONE
019250 PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
019300 PERFORM COL-TITLES-C VARYING A FROM 27 BY 1 UNTIL A > 33. ANAL-ONE
019350 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
019400 PERFORM OUTPT-TWO-DUMP-C VARYING B FROM 34 BY 1 UNTIL B > 56. ANAL-ONE
019450 PERFORM OUTPT-THREE-PRP. ANAL-ONE
019500 PERFORM COL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
019550 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
019600 PERFORM OUTPT-TWO-DUMP-A VARYING B FROM 57 BY 1 UNTIL B > 66. ANAL-ONE
019650 PERFORM COL-TITLES-B VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-ONE
019700 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
019750 PERFORM OUTPT-TWO-DUMP-B VARYING B FROM 57 BY 1 UNTIL B > 66. ANAL-ONE
019800 PERFORM OUTPT-THREE-PRP. ANAL-ONE
019850 PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
019900 PERFORM COL-TITLES-C VARYING A FROM 27 BY 1 UNTIL A > 33. ANAL-ONE
020000 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
020050 PERFORM OUTPT-TWO-DUMP-C VARYING B FROM 57 BY 1 UNTIL B > 66. ANAL-ONE
020100 PERFORM VC-DUMP. ANAL-ONE
020150 PERFORM YULE-PRP VARYING A FROM 1 BY 1 UNTIL A > 33. ANAL-ONE
020200 COMPUTE YULES-K ROUNDED = (10000 * (S - ILF-TOTAL)) / ANAL-ONE
020250 (ILF-TOTAL ** 2). ANAL-ONE
020300 WRITE PRINT-LINE FROM OUTPT-FOUR-E BEFORE 4. ANAL-ONE
020350 WRITE PRINT-LINE FROM OUTPT-FOUR-F BEFORE 2. ANAL-ONE
020400 WRITE PRINT-LINE FROM OUTPT-FOUR-G BEFORE 2. ANAL-ONE
020450 PERFORM KS-CALC VARYING A FROM 1 BY 1 UNTIL A > 33. ANAL-ONE
020500 PERFORM JOINT-PRB. ANAL-ONE
INITIALIZE-RT. ANAL-ONE
020550 MOVE LANG-KEY TO LANG-ID. ANAL-ONE
020600 MOVE ZEROS TO VC-TOTAL, A, B, N, F-CONS-CONS, F-VOW-CONS, ANAL-ONE
020650 MOVE F-CONS-VOW, F-VOW-VOW, FOL-FREQ-TOTAL, S, ILF-TOTAL, ANAL-ONE
020700 CRF-TOTAL, CARD-NR, L, M. ANAL-ONE
020750 PERFORM ZEROING-ONE VARYING A FROM 1 BY 1 UNTIL A > 33. ANAL-ONE
020800 PERFORM ZEROING-TWO VARYING A FROM 1 BY 1 UNTIL A > 33. ANAL-ONE
020850 PERFORM AFTER ZEROS FROM 1 BY 1 UNTIL B > 66. ANAL-ONE
020900 MOVE LANGUAGE TO LANG-DESIG, OUTPT-FOUR-LANG, ANAL-ONE
020950 MOVE OUTPT-FOUR-KS-LANG. ANAL-ONE
021000 READ POPULATION AT END GO TO DUMP. ANAL-ONE
021050 IND-TOTAL-CALC. ANAL-ONE
021100 COMPUTE ILF-TOTAL = ILF-TOTAL + ILF (A). ANAL-ONE
021150 IND-PRB-CALC. ANAL-ONE
021200 IF ILF (A) > 0 COMPUTE ILP (A) ROUNDED = ANAL-ONE
021250 (ILF (A) / ILF-TOTAL); ELSE MOVE ZEROS TO ILP (A). ANAL-ONE
021300 OUTPT-ONE-PRP. ANAL-ONE
021350 MOVE SPACES TO PRINT-LINE. ANAL-ONE
021400

```



```

021450 WRITE LINE FROM OUTPT-CNE-A AFTER TOP-OF-PAGE. ANAL-ONE
021500 PRINT-LINE FROM OUTPT-CNE-B BEFORE 1. ANAL-ONE
021550 PRINT-LINE FROM LANG-LINE AFTER 1. ANAL-ONE
021600 PRINT-LINE FROM LANG-UNDERLINE BEFORE 1. ANAL-ONE
021650 PRINT-LINE FROM OUTPT-CNE-C AFTER 2. ANAL-ONE
021700 PRINT-LINE FROM OUTPT-CNE-D BEFORE 2. ANAL-ONE
021750 CUTPT-ONE-PRINT: TO LTR-ONE. ANAL-ONE
021800 MOVE LTR (A) TO FREQ-ONE. ANAL-ONE
021850 MOVE ILP (A) TO PROB-ONE. ANAL-ONE
021900 MOVE ILP (A) TO PROB-ONE. ANAL-ONE
021950 WRITE PRINT-LINE FROM OUTPT-CNE-E BEFORE 2. ANAL-ONE
022000 DEPN-TWO-CALC. TO FOL-FREQ-TOTAL. ANAL-ONE
022050 MOVE ZEROS. TO FOL-FREQ-TOTAL. ANAL-ONE
022100 PERFORM DEPN-TWO-SUM VARYING B FROM 1 BY 1 UNTIL B > 33. ANAL-ONE
022150 PERFORM DEPN-TWO-PROB VARYING B FROM 1 BY 1 UNTIL B > 33. ANAL-ONE
022200 DEPN-TWO-SUM. FOL-FREQ-TOTAL = FOL-FREQ-TOTAL + FOL-FREQ (A, B). ANAL-ONE
022250 COMPUTE FOL-FREQ (A, B) > 0 COMPUTE FOL-PROB (A, B) ROUNDED = ANAL-ONE
022300 IF FOL-FREQ (A, B) / FOL-FREQ-TOTAL, ELSE MOVE ZEROS ANAL-ONE
022350 {FOL-FREQ (A, B) / FOL-FREQ-TOTAL}, ELSE MOVE ZEROS ANAL-ONE
022400 TO FOL-PROB (A, B). ANAL-ONE
022450 OUTPT-TWO-SPACES. TO PRINT-LINE. ANAL-ONE
022500 MOVE SPACES. TO PRINT-LINE. ANAL-ONE
022550 WRITE PRINT-LINE FROM OUTPT-TWO-A AFTER TOP-OF-PAGE. ANAL-ONE
022600 WRITE PRINT-LINE FROM OUTPT-TWO-THREE-B BEFORE 1. ANAL-ONE
022650 WRITE PRINT-LINE FROM LANG-LINE BEFORE 0. ANAL-ONE
022700 WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2. ANAL-ONE
022750 WRITE PRINT-LINE FROM FIRST-LTR-OUTPT BEFORE 2. ANAL-ONE
022800 COL-TITLE-A. (A) TO FIRST-LTR (A). ANAL-ONE
022850 MOVE LTR (A) TO FIRST-LTR (A). ANAL-ONE
022900 CUTPT-TWO-DUMP-A. A VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
022950 PERFORM DUMP-A. FROM DEPN-FREQ-LINE BEFORE 1. ANAL-ONE
023000 WRITE PRINT-LINE FROM DEPN-FREQ-LINE BEFORE 1. ANAL-ONE
023050 PERFORM DUMP-B VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
023100 PERFORM DUMP-B. FROM DEPN-PROB-LINE BEFORE 2. ANAL-ONE
023150 WRITE PRINT-LINE FROM DEPN-PROB-LINE BEFORE 2. ANAL-ONE
023200 CUTPT-TWO-DUMP-B. A VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-ONE
023250 PERFORM DUMP-B. FROM DEPN-FREQ-LINE BEFORE 1. ANAL-ONE
023300 WRITE PRINT-LINE FROM DEPN-FREQ-LINE BEFORE 1. ANAL-ONE
023350 PERFORM DUMP-B VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-ONE
023400 WRITE PRINT-LINE FROM DEPN-PROB-LINE BEFORE 2. ANAL-ONE
023450 CUTPT-TWO-DUMP-C. A VARYING A FROM 27 BY 1 UNTIL A > 33. ANAL-ONE
023500 PERFORM DUMP-C. FROM DEPN-FREQ-LINE BEFORE 1. ANAL-ONE
023550 WRITE PRINT-LINE FROM DEPN-FREQ-LINE BEFORE 1. ANAL-ONE
023600 PERFORM DUMP-B VARYING A FROM 27 BY 1 UNTIL A > 33. ANAL-ONE
023650 WRITE PRINT-LINE FROM DEPN-PROB-LINE BEFORE 2. ANAL-ONE
023700 COL-TITLE-B. L = A - 13. ANAL-ONE
023750 MOVE LTR (A) TO FIRST-LTR (L). ANAL-ONE
023800

```



```

023850 CCL-TITLES-C.
023900 COMPUTE L = A - 26.
023950 MOVE LTR (A) TO FIRST-LTR (L).
024000 DEPN-TWO-THREE-CALC.
024050 MOVE ZEROS TO FOL-FREQ-TOTAL.
024100 PERFORM DEPN-TWO-SUM VARYING B FROM 34 BY 1 UNTIL B > 66.
024150 PERFORM DEPN-TWO-PRCB VARYING B FROM 34 BY 1 UNTIL B > 66.
024200 CUTPT-THREE-PREP.
024250 MOVE SPACES TO PRINT-LINE.
024300 WRITE PRINT-LINE FROM OUTPT-THREE-A AFTER TOP-OF-PAGE.
024350 WRITE PRINT-LINE FROM OUTPT-TWO-THREE-B BEFORE 1.
024400 WRITE PRINT-LINE FROM LANG-LINE BEFORE 0.
024450 WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2.
024500 WRITE PRINT-LINE FROM FIRST-LTR-OUTPT BEFORE 2.
024550 VC-DUMP.
024600 COMPUTE VC-TOTAL = F-CONS-CONS + F-VOW-CONS + F-CONS-VOW
024650 + F-VOW-VOW.
024700 IF F-VOW-VOW > 0 COMPUTE P-VOW-VOW ROUNDED = F-VOW-VOW /
024750 VC-TOTAL ELSE MOVE ZEROS TO P-VOW-VOW.
024800 IF F-CONS-VOW > 0 COMPUTE P-CONS-VOW ROUNDED = F-CONS-VOW /
024850 VC-TOTAL ELSE MOVE ZEROS TO P-CONS-VOW.
024900 IF F-VOW-CONS > 0 COMPUTE P-VOW-CONS ROUNDED = F-VOW-CONS /
024950 VC-TOTAL ELSE MOVE ZEROS TO P-VOW-CONS.
025000 IF F-CONS-CONS > 0 COMPUTE P-CONS-CONS ROUNDED = F-CONS-CONS /
025050 VC-TOTAL ELSE MOVE ZEROS TO P-CONS-CONS.
025100 WRITE PRINT-LINE FROM LANG-LINE AFTER TOP-OF-PAGE.
025150 WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2.
025200 WRITE PRINT-LINE FROM OUTPT-FOUR-A BEFORE 2.
025250 WRITE PRINT-LINE FROM OUTPT-FOUR-B BEFORE 2.
025300 WRITE PRINT-LINE FROM OUTPT-FOUR-C BEFORE 2.
025350 WRITE PRINT-LINE FROM OUTPT-FOUR-D BEFORE 3.
025400 YULE-PREP.
025450 COMPUTE S = S + (ILF (A) ** 2).
025500 KS-CALC.
025550 COMPUTE CRF-TOTAL = CRF-TOTAL + ILP (A).
025600 MOVE LTR (A) TO KS-LTR.
025650 MOVE CRF-TOTAL TO KS-CRF.
025700 WRITE PRINT-LINE FROM KS-LINE BEFORE 1.
025750 E-1. MOVE 014300 TO ER-STEP.
025800 GO TO DUMP.
025850 E-2. MOVE 014901 TO ER-STEP.
025900 GO TO DUMP.
025950 E-3. MOVE 015400 TO ER-STEP.
026000 GO TO DUMP.
026050 E-4. MOVE 015400 TO ER-STEP.
026100 GO TO DUMP.
026150
026200

```


026250	MOVE 015450 TO ER-STEP.	ANAL-ONE
026300	GO TO DUMP.	ANAL-ONE
026350	E-5.	ANAL-ONE
026400	MOVE 015550 TO ER-STEP.	ANAL-ONE
026450	GO TO DUMP.	ANAL-ONE
026500	E-6.	ANAL-ONE
026550	MOVE 015650 TO ER-STEP.	ANAL-ONE
026600	GO TO DUMP.	ANAL-ONE
026650	E-7.	ANAL-ONE
026700	MOVE 015750 TO ER-STEP.	ANAL-ONE
026750	GO TO DUMP.	ANAL-ONE
026800	E-8.	ANAL-ONE
026850	MOVE 016000 TO ER-STEP.	ANAL-ONE
026900	GO TO DUMP.	ANAL-ONE
026950	E-9.	ANAL-ONE
027000	MOVE 017300 TO ER-STEP.	ANAL-ONE
027050	GO TO DUMP.	ANAL-ONE
027100	E-10.	ANAL-ONE
027150	MOVE 017500 TO ER-STEP.	ANAL-ONE
027200	DUMP.	ANAL-ONE
027250	MOVE CARD-NR TO ER-CARD.	ANAL-ONE
027300	WRITE PRINT-LINE FROM ERROR-LINE AFTER 2.	ANAL-ONE
027350	GO TO EQJ.	ANAL-ONE
027400	DUMP-A.	ANAL-ONE
027450	IF B > 33 COMPUTE L = B - 33, MOVE LTR (L) TO FREQ-ROW-LTR,	ANAL-ONE
027500	ELSE MOVE LTR (B) TO FREQ-ROW-LTR.	ANAL-ONE
027550	IF A < 14 MOVE FOL-FREQ (A, B) TO DEPN-FREQ (A).	ANAL-ONE
027600	IF A > 13 AND A < 27 COMPUTE M = A - 13, MOVE FCL-FREQ (A, B)	ANAL-ONE
027650	TO DEPN-FREQ (M).	ANAL-ONE
027700	IF A > 26 COMPUTE M = A - 26, MOVE FOL-FREQ (A, B)	ANAL-ONE
027750	TO DEPN-FREQ (M).	ANAL-ONE
027800	DUMP-B.	ANAL-ONE
027850	IF B > 33 COMPUTE L = B - 33, MOVE LTR (L) TO PROB-ROW-LTR,	ANAL-ONE
027900	ELSE MOVE LTR (B) TO PROB-ROW-LTR.	ANAL-ONE
027950	IF A < 14 MOVE FOL-PROB (A, B) TO DEPN-PROB (A).	ANAL-ONE
028000	IF A > 13 AND A < 27 COMPUTE M = A - 13, MOVE FCL-PROB (A, B)	ANAL-ONE
028050	TO DEPN-PROB (M).	ANAL-ONE
028100	IF A > 26 COMPUTE M = A - 26, MOVE FOL-PROB (A, B)	ANAL-ONE
028150	TO DEPN-PROB (M).	ANAL-ONE
028200	ZEROING-ONE.	ANAL-ONE
028250	MOVE ZEROS TO ILF (A).	ANAL-ONE
028300	MOVE ZEROS TO ILP (A).	ANAL-ONE
028350	FLUSHING.	ANAL-ONE
028400	MOVE SPACES TO FIRST-LTR (A).	ANAL-ONE
028450	MOVE ZEROS TO DEPN-FREQ (A).	ANAL-ONE
028500	MOVE ZEROS TO DEPN-PROB (A).	ANAL-ONE
028550	ZEROING-TWO.	ANAL-ONE
028600	MOVE ZEROS TO FOL-FREQ (A, B).	ANAL-ONE


```

028650 MOVE ZEROS TO FOL-PROB (A, B).
028700 ADER.
028750 COMPUTE B = B + 33.
028800 ADD 1 TO FOL-FREQ (A, B) ON SIZE ERROR GO TO E-8.
028850 COMPUTE B = B - 33.
028900 DAT-SHIFT.
028950 COMPUTE Y = X + 2 * (Y).
029000 MOVE CB (X) TO CC (Y).
029050 JOINT-PROB.
029100 MOVE ZEROS TO FOL-FREQ-TOTAL.
029150 PERFORM DEPN-TWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 33
029200 AFTER B FROM 1 BY 1 UNTIL B > 33.
029250 PERFORM DEPN-TWO-PROB VARYING A FROM 1 BY 1 UNTIL A > 33
029300 AFTER B FROM 1 BY 1 UNTIL B > 33.
029350 PERFORM OUTPT-FIVE-PREP.
029400 PERFORM CCL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13.
029450 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.
029500 PERFORM OUTPT-DUMP-D VARYING B FROM 1 BY 1 UNTIL B > 33.
029550 PERFORM OUTPT-FIVE-PREP.
029600 PERFORM COL-TITLES-B VARYING A FROM 14 BY 1 UNTIL A > 26.
029650 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.
029700 PERFORM OUTPT-DUMP-E VARYING B FROM 1 BY 1 UNTIL B > 33.
029750 PERFORM OUTPT-FIVE-PREP.
029800 PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13.
029850 PERFORM COL-TITLES-C VARYING A FROM 27 BY 1 UNTIL A > 33.
029900 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.
029950 PERFORM OUTPT-DUMP-F VARYING B FROM 1 BY 1 UNTIL B > 33.
030000 MOVE ZEROS TO S.
030050 PERFORM YULES-K-PREP VARYING A FROM 1 BY 1 UNTIL A > 33
030100 AFTER B FROM 1 BY 1 UNTIL B > 33.
030150 COMPUTE YULES-K-2 ROUNDED = (10000 * (S - FOL-FREQ-TOTAL)) /
030200 (FOL-FREQ-TOTAL ** 2).
030250 WRITE PRINT-LINE FROM OUTPT-FIVE-B BEFORE 1.
030300 MOVE ZEROS TO CRF-TOTAL, WS-SEQ-NR, L.
030350 MOVE B TO TEST.
030400 PERFORM CARD-DUMP VARYING A FROM 1 BY 1 UNTIL A > 33
030450 AFTER B FROM 1 BY 1 UNTIL B > 33.
030500 MOVE ZEROS TO FOL-FREQ-TOTAL.
030550 PERFORM DEPN-TWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 33
030600 AFTER B FROM 34 BY 1 UNTIL B > 66.
030650 PERFORM DEPN-TWO-PROB VARYING A FROM 1 BY 1 UNTIL A > 33
030700 AFTER B FROM 34 BY 1 UNTIL B > 66.
030750 PERFORM OUTPT-SIX-PREP.
030800 PERFORM COL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13.
030850 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.
030900 PERFORM OUTPT-DUMP-D VARYING B FROM 34 BY 1 UNTIL B > 66.
030950 PERFORM OUTPT-SIX-PREP.
031000 PERFORM COL-TITLES-B VARYING A FROM 14 BY 1 UNTIL A > 26.

```



```

031050 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
031100 PERFORM OUTPT-DUMP-E VARYING B FROM 34 BY 1 UNTIL B > 66. ANAL-ONE
031150 PERFORM OUTPT-SIX-PRP. ANAL-ONE
031200 PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
031250 PERFORM COL-TITLES-C VARYING A FROM 27 BY 1 UNTIL A > 33. ANAL-ONE
031300 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-ONE
031350 PERFORM OUTPT-DUMP-F VARYING B FROM 34 BY 1 UNTIL B > 66. ANAL-ONE
031400 MOVE ZEROS TO S. ANAL-ONE
031450 PERFORM YULES-K-PRP VARYING A FROM 1 BY 1 UNTIL A > 33 ANAL-ONE
031500 AFTER B FROM 34 BY 1 UNTIL B > 66. ANAL-ONE
031550 COMPUTE YULES-K-3 ROUNDED = (10000 * (S - FOL-FREQ-TOTAL)) / ANAL-ONE
031600 (FOL-FREQ-TOTAL ** 2). ANAL-ONE
031650 WRITE PRINT-LINE FROM OUTPT-SIX-B BEFORE 1. ANAL-ONE
031700 MOVE ZEROS TO CRF-TOTAL, WS-SEQ-NR, L. ANAL-ONE
031750 MOVE ,C, TO TEST. ANAL-ONE
031800 PERFORM CARD-DUMP VARYING A FROM 1 BY 1 UNTIL A > 33 ANAL-ONE
031850 AFTER B FROM 34 BY 1 UNTIL B > 66. ANAL-ONE
031900 CUTPT-FIVE-PRP. ANAL-ONE
031950 WRITE PRINT-LINE FROM OUTPT-FIVE-A AFTER TOP-OF-PAGE. ANAL-ONE
032000 WRITE PRINT-LINE FROM OUTPT-TWO-THREE-B BEFORE 1. ANAL-ONE
032050 WRITE PRINT-LINE FROM LANG-LINE BEFORE 0. ANAL-ONE
032100 WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2. ANAL-ONE
032150 WRITE PRINT-LINE FROM FIRST-LTR-OUTPT BEFORE 2. ANAL-ONE
032200 OUTPT-DUMP-D. ANAL-ONE
032250 PERFORM DUMP-B VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-ONE
032300 WRITE PRINT-LINE FROM DEPN-PROB-LINE BEFORE 2. ANAL-ONE
032350 CUTPT-DUMP-E. ANAL-ONE
032400 PERFORM DUMP-B VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-ONE
032450 WRITE PRINT-LINE FROM DEPN-PROB-LINE BEFORE 2. ANAL-ONE
032500 CUTPT-DUMP-F. ANAL-ONE
032550 PERFORM DUMP-B VARYING A FROM 27 BY 1 UNTIL A > 33. ANAL-ONE
032600 WRITE PRINT-LINE FROM DEPN-PROB-LINE BEFORE 2. ANAL-ONE
032650 CUTPT-SIX-PRP. ANAL-ONE
032700 WRITE PRINT-LINE FROM OUTPT-SIX-A AFTER TOP-OF-PAGE. ANAL-ONE
032750 WRITE PRINT-LINE FROM OUTPT-TWO-THREE-B BEFORE 1. ANAL-ONE
032800 WRITE PRINT-LINE FROM LANG-LINE BEFORE 0. ANAL-ONE
032850 WRITE PRINT-LINE FROM LANG-UNDERLINE BEFORE 2. ANAL-ONE
032900 WRITE PRINT-LINE FROM FIRST-LTR-OUTPT BEFORE 2. ANAL-ONE
032950 YULES-K-PRP. ANAL-ONE
033000 COMPUTE S = S + (FOL-FREQ (A, B) ** 2). ANAL-ONE
033050 CARD-DUMP. ANAL-ONE
033100 COMPUTE CRF-TOTAL = CRF-TOTAL + FOL-PRCB (A, B). ANAL-ONE
033150 ADD 1 TO CRF-TOTAL TO CRF-OUT (L). ANAL-ONE
033200 MOVE 1 TO WS-SEQ-NR ANAL-ONE
033250 IF L = 11 MOVE WS-SEQ-NR TO SEQ-NR ANAL-ONE
033300 MOVE ZEROS TO L MOVE CARD-OUT FROM CARD-FILL. ANAL-ONE
033350 ANAL-ONE
033400 ANAL-ONE

```


APPENDIX H. POPULATION ANALYSIS PROGRAM (V-2)

```

0000050 IDENTIFICATION DIVISION.
0000100 PROGRAM-ID. V2-LANGUAGE-ANALYSIS.
0000150 AUTHOR. LCOL. MORTON D. RAU, USN.
0000200 REMARKS. THIS PROGRAM IS PART OF MY MASTER OF SCIENCE IN
0000250 MANAGERIAL THESIS AND IS DESIGNED TO CALCULATE VARIOUS
0000300 STATISTICAL CHARACTERISTICS FOR LETTERS APPEARING IN
0000350 LANGUAGES WRITTEN IN THE MODERN LATIN ALPHABET.
0000400 VERSION TWO IS AN EXPANSION DESIGNED TO INCLUDE THE SPACE
0000450 RATE IN THESE LANGUAGES AND TO INCLUDE THE INTERDEPENDENCY
0000500 OF LETTERS IN ADJOINING WORDS.
0000550 ENVIRONMENT DIVISION.
0000600 CONFIGURATION SECTION.
0000650 SOURCE-COMPUTER. IBM-360-67.
0000700 OBJECT-COMPUTER. IBM-360-67.
0000750 SPECIAL-NAMES.
0000800 CO1 IS TOP-OF-PAGE.
0000850 INPUT-CONTROL.
0000900 FILE-CONTROL.
0000950 SELECT POPULATION ASSIGN TO UR-S-IN1.
0001000 SELECT LANG-ANAL ASSIGN TO UR-S-OUT1.
0001050 SELECT CARD-FILE ASSIGN TO UR-S-OUT2.
0001100 DATA DIVISION.
0001150 FILE SECTION.
0001200 FD POPULATION.
0001250 LABEL RECORDS ARE OMITTED
0001300 BLOCK CONTAINS 5 RECORDS
0001350 DATA RECORD IS POP-CARD, LANG-CARD.
0001400 POP-CARD.
0001450 05 FILLER
0001500 05 SAMPLE-FLD.
0001550 05 IO CC
0001600 05 FILLER
0001650 05 LANG-CARD-TYPE
0001700 05 88 NEW-LANG VALUE 'L'.
0001750 05 FILLER
0001800 05 LANG
0001850 05 FILLER
0001900 05 FILLER
0001950 05 LANG-KEY
0002000 05 LANG-ANAL
0002050 LABEL RECORDS ARE OMITTED
0002100 BLOCK CONTAINS 5 RECORDS
0002150 DATA RECORD IS PRINT-LINE.
0002200 01 PRINT-LINE

```


004650	05	FILLER	PIC X	VALUE 'Q'.	ANAL-TWO
004700	05	FILLER	PIC X	VALUE 'R'.	ANAL-TWO
004750	05	FILLER	PIC X	VALUE 'S'.	ANAL-TWO
004800	05	FILLER	PIC X	VALUE 'T'.	ANAL-TWO
004850	05	FILLER	PIC X	VALUE 'V'.	ANAL-TWO
004900	05	FILLER	PIC X	VALUE 'W'.	ANAL-TWO
004950	05	FILLER	PIC X	VALUE 'X'.	ANAL-TWO
005000	05	FILLER	PIC X	VALUE 'Y'.	ANAL-TWO
005050	05	FILLER	PIC X	VALUE 'Z'.	ANAL-TWO
005100	05	FILLER	PIC X	VALUE '6'.	ANAL-TWO
005150	05	FILLER	PIC X	VALUE '7'.	ANAL-TWO
005200	05	FILLER	PIC X	VALUE '1'.	ANAL-TWO
005250	01	LTR-TABLE	REDEFINES LTR-TABLE-VALUES.		ANAL-TWO
005300	05	LTR	PIC X,	OCCURS 34 TIMES.	ANAL-TWO
005350	01	IND-FREQ-TABLE.			ANAL-TWO
005400	05	ILF-PIC 9(6),	COMP, SYNC,	OCCURS 34 TIMES.	ANAL-TWO
005450	01	IND-PROB-TABLE.			ANAL-TWO
005500	05	ILP-PIC 9V9(5),	COMP, SYNC,	OCCURS 34 TIMES.	ANAL-TWO
005550	01	DEPENDENCY-FREQ-TABLE.			ANAL-TWO
005600	05	FST-LTR	OCCURS 34 TIMES.		ANAL-TWO
005650	01	IO FOL-FREQ	PIC 9(6),	COMP, SYNC,	OCCURS 68 TIMES.
005700	05	DEPENDENCY-PROB-TABLE.			ANAL-TWO
005750	05	GIVEN-LTR	OCCURS 34 TIMES.		ANAL-TWO
005800	01	IO FOL-PROB	PIC 9V9(5),	COMP, SYNC,	OCCURS 68 TIMES.
005850	01	OUTPT-ONE-A.			ANAL-TWO
005900	05	FILLER	PIC X(43)	VALUE SPACES.	ANAL-TWO
005950	05	FILLER	PIC X(48)	VALUE 'FREQUENCY/INDEPE	ANAL-TWO
006000	-	IDENT PROBABILITY OF EACH LETTER.			ANAL-TWO
006050	01	FILLER	PIC X(42)	VALUE SPACES.	ANAL-TWO
006100	05	LANG-LINE.			ANAL-TWO
006150	05	FILLER	PIC X(54)	VALUE SPACES.	ANAL-TWO
006200	05	FILLER	PIC X(11)	VALUE 'LANGUAGE:'. .	ANAL-TWO
006250	05	LANG-DESIG	PIC X(30)	VALUE SPACES.	ANAL-TWO
006300	05	FILLER	PIC X(38)	VALUE SPACES.	ANAL-TWO
006350	01	LANG-UNDERLINE.			ANAL-TWO
006400	05	FILLER	PIC X(65)	VALUE SPACES.	ANAL-TWO
006450	05	FILLER	PIC X(30)	VALUE ALL '!'.	ANAL-TWO
006500	05	FILLER	PIC X(38)	VALUE SPACES.	ANAL-TWO
006550	01	OUTPT-ONE-B.			ANAL-TWO
006600	05	FILLER	PIC X(43)	VALUE SPACES.	ANAL-TWO
006650	05	FILLER	PIC X(48)	VALUE ALL '!'.	ANAL-TWO
006700	05	FILLER	PIC X(42)	VALUE SPACES.	ANAL-TWO
006750	01	OUTPT-ONE-C.			ANAL-TWO
006800	05	FILLER	PIC X(37)	VALUE SPACES.	ANAL-TWO
006850	05	FILLER	PIC X(25)	VALUE 'LETTER'.	ANAL-TWO
006900	05	FILLER	PIC X(25)	VALUE 'FREQUENCY'.	ANAL-TWO
006950	05	FILLER	PIC X(46)	VALUE 'PROBABILITY'.	ANAL-TWO
007000	01	OUTPT-ONE-D.			ANAL-TWO

014250	MOVE 3 TO N.	ANAL-TWO
014300	BRAVO.	ANAL-TWO
014350	IF CL (N) = SPACE GO TO CK-ONE.	ANAL-TWO
014400	PERFORM FIRST-LETTER-CHECK THRU FIRST-EXIT.	ANAL-TWO
014450	IF MATCH = ZERO GO TO CK-ONE.	ANAL-TWO
014500	ADD 1 TO ILF (A) ON SIZE ERROR GO TO E-1.	ANAL-TWO
014550	COMPUTE M = N + 1. SHIFT-TWO.	ANAL-TWO
014600	IF M = 73 PERFORM GO TO CK-TWO.	ANAL-TWO
014650	IF CL (M) = SPACE GO TO CK-TWO.	ANAL-TWO
014700	PERFORM SECOND-LETTER-CHECK THRU SECOND-EXIT.	ANAL-TWO
014750	IF MATCH = ZERO GO TO CK-TWO.	ANAL-TWO
014800	CHARLIE.	ANAL-TWO
014850	ADD 1 TO FOL-FREQ (A, B) ON SIZE ERROR GO TO E-2.	ANAL-TWO
014900	PERFORM VC-TALLEY THRU VC-EXIT.	ANAL-TWO
014950	COMPUTE L = N + 2.	ANAL-TWO
015000	IF L = 73 AND CL (73) = ' ' PERFORM SHIFT-THREE.	ANAL-TWO
015050	IF L = 73 AND CL (73) NOT = ' ' PERFORM ALT-THREE.	ANAL-TWO
015100	IF CL (L) = SPACE GO TO CK-THREE.	ANAL-TWO
015150	PERFORM THIRD-LETTER-CHECK THRU THIRD-EXIT.	ANAL-TWO
015200	IF MATCH = ZERO GO TO CK-THREE.	ANAL-TWO
015250	DELTA.	ANAL-TWO
015300	COMPUTE B = B + 34.	ANAL-TWO
015350	ADD 1 TO FOL-FREQ (A, B) ON SIZE ERROR GO TO E-3.	ANAL-TWO
015400	ECHO.	ANAL-TWO
015450	ADD 1 TO N.	ANAL-TWO
015500	GO TO BRAVO.	ANAL-TWO
015550	CK-ONE.	ANAL-TWO
015600	ADD 1 TO ILF (34) ON SIZE ERROR GO TO E-4.	ANAL-TWO
015650	ADD 1 TO N.	ANAL-TWO
015700	MOVE 34 TO A.	ANAL-TWO
015750	CK-ONE-ALPHA.	ANAL-TWO
015800	IF N = 73 PERFORM SHIFT-ONE.	ANAL-TWO
015850	IF CL (N) = SPACE GO TO CK-ONE-BRAVO.	ANAL-TWO
015900	COMPUTE N = N - 1.	ANAL-TWO
015950	PERFORM SECOND-LETTER-CHECK THRU SECOND-EXIT.	ANAL-TWO
016000	IF MATCH = 1 GO TO CHARLIE, ELSE ADD 1 TO N.	ANAL-TWO
016050	CK-ONE-BRAVO.	ANAL-TWO
016100	ADD 1 TO N.	ANAL-TWO
016150	GO TO CK-ONE-ALPHA.	ANAL-TWO
016200	CK-TWO.	ANAL-TWO
016250	ADD 1 TO FOL-FREQ (A, 34) ON SIZE ERROR GO TO E-5.	ANAL-TWO
016300	ADD 1 TO N.	ANAL-TWO
016350	CK-TWO-ALPHA.	ANAL-TWO
016400	COMPUTE M = N + 1. SHIFT-TWO.	ANAL-TWO
016450	IF M = 73 PERFORM GO TO CK-TWO-BRAVO.	ANAL-TWO
016500	IF CL (M) = SPACE GO TO CK-TWO-BRAVO.	ANAL-TWO
016550	COMPUTE N = N - 1.	ANAL-TWO
016600	PERFORM THIRD-LETTER-CHECK THRU THIRD-EXIT.	ANAL-TWO

016650	IF MATCH = 1 GO TO DELTA, ELSE ADD 1 TO N.	ANAL-TWO
016700	CK-TWO-BRavo.	ANAL-TWO
016750	ADD 1 TO N.	ANAL-TWO
016800	GO TO CK-TWO-ALPHA.	ANAL-TWO
016850	CK-THREE.	ANAL-TWO
016900	ADD 1 TO FOL-FREQ (A, 68) ON SIZE ERROR GO TO E-6.	ANAL-TWO
016950	GO TO ECHO.	ANAL-TWO
017000	SHIFT-ONE.	ANAL-TWO
017050	PERFORM READ-A-CARD.	ANAL-TWO
017100	MOVE SPACES TO CL (1), CL (2).	ANAL-TWO
017150	MOVE 3 TO N.	ANAL-TWO
017200	SHIFT-TWO.	ANAL-TWO
017250	MOVE CL (72) TO CL (2).	ANAL-TWO
017300	PERFORM READ-A-CARD.	ANAL-TWO
017350	MOVE 2 TO N.	ANAL-TWO
017400	MOVE 3 TO M.	ANAL-TWO
017450	SHIFT-THREE.	ANAL-TWO
017500	MOVE CL (72) TO CL (2).	ANAL-TWO
017550	PERFORM READ-A-CARD.	ANAL-TWO
017600	MOVE 1 TO N.	ANAL-TWO
017650	MOVE 3 TO L.	ANAL-TWO
017700	ALT-THREE.	ANAL-TWO
017750	MOVE CL (72) TO CL (1).	ANAL-TWO
017800	MOVE SPACES TO CL (2).	ANAL-TWO
017850	PERFORM READ-A-CARD.	ANAL-TWO
017900	MOVE 0 TO N.	ANAL-TWO
017950	MOVE 2 TO L.	ANAL-TWO
018000	READ-A-CARD.	ANAL-TWO
018050	READ POPULATION AT END GO TO EOJ.	ANAL-TWO
018100	IF NEW-LANG GO TO FLUSH.	ANAL-TWO
018150	ADD 1 TO CARD-NR ON SIZE ERROR GO TO E-7.	ANAL-TWO
018200	PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.	ANAL-TWO
018250	DAT-SHIFT.	ANAL-TWO
018300	PERFORM COMPUTE Y = X + 2.	ANAL-TWO
018350	MOVE CC (X) TO CL (Y).	ANAL-TWO
018400	FLUSH.	ANAL-TWO
018450	IF CARD-NR NOT = ZERO PERFORM CLOSE-OUT-RT.	ANAL-TWO
018500	PERFORM INITIALIZE-RT.	ANAL-TWO
018550	GO TO ALPHA.	ANAL-TWO
018600	EOJ.	ANAL-TWO
018650	PERFORM CLOSE-OUT-RT.	ANAL-TWO
018700	PERFORM LOOP-RT 5 TIMES.	ANAL-TWO
018750	MOVE SPACES TO PRINT-LINE.	ANAL-TWO
018800	WRITE PRINT-LINE AFTER TOP-OF-PAGE.	ANAL-TWO
018850	CLOSE POPULATION, LANG-ANAL, CARD-FILE.	ANAL-TWO
018900	STOP RUN.	ANAL-TWO
018950	LCOP-RT.	ANAL-TWO
019000	MOVE SPACES TO PRINT-LINE.	ANAL-TWO

019050	WRITE PRINT-LINE AFTER 1.	ANAL-TWO
019100	FIRST-LETTER-CHECK.	ANAL-TWO
019150	MOVE 1 TO A.	ANAL-TWO
019200	MOVE ZERO TO MATCH.	ANAL-TWO
019250	FLC-ALPHA.	ANAL-TWO
019300	IF CL (N) = LTR (A) GO TO FLC-BRAVO.	ANAL-TWO
019350	IF A = 33 GO TO FIRST-EXIT.	ANAL-TWO
019400	ADD 1 TO A.	ANAL-TWO
019450	GO TO FLC-ALPHA.	ANAL-TWO
019500	FLC-BRAVO.	ANAL-TWO
019550	MOVE 1 TO MATCH.	ANAL-TWO
019600	FIRST-EXIT.	ANAL-TWO
019650	SECOND-LETTER-CHECK.	ANAL-TWO
019700	MOVE 1 TO B.	ANAL-TWO
019750	MOVE ZERO TO MATCH.	ANAL-TWO
019800	COMPUTE M = N + 1.	ANAL-TWO
019850	SLC-ALPHA.	ANAL-TWO
019900	IF CL (M) = LTR (B) GO TO SLC-BRAVO.	ANAL-TWO
019950	IF B = 33 GO TO SECOND-EXIT.	ANAL-TWO
020000	ADD 1 TO B.	ANAL-TWO
020050	GO TO SLC-ALPHA.	ANAL-TWO
020100	SLC-BRAVO.	ANAL-TWO
020150	MOVE 1 TO MATCH.	ANAL-TWO
020200	SECOND-EXIT.	ANAL-TWO
020250	THIRD-LETTER-CHECK.	ANAL-TWO
020300	MOVE 1 TO B.	ANAL-TWO
020350	MOVE ZERO TO MATCH.	ANAL-TWO
020400	COMPUTE L = N + 2.	ANAL-TWO
020450	TLC-ALPHA.	ANAL-TWO
020500	IF CL (L) = LTR (B) GO TO TLC-BRAVO.	ANAL-TWO
020550	IF B = 33 GO TO THIRD-EXIT.	ANAL-TWO
020600	ADD 1 TO B.	ANAL-TWO
020650	GO TO TLC-ALPHA.	ANAL-TWO
020700	TLC-BRAVO.	ANAL-TWO
020750	MOVE 1 TO MATCH.	ANAL-TWO
020800	THIRD-EXIT.	ANAL-TWO
020850	VC-TALLEY.	ANAL-TWO
020900	IF A = 34 GO TO VC-EXIT.	ANAL-TWO
020950	IF A < 11 AND B < 11 ADD 1 TO F-VOW-VOW SIZE ERROR GO TO E-8.	ANAL-TWO
021000	IF A < 11 AND B NOT < 11 ADD 1 TO F-CONS-VOW ON SIZE	ANAL-TWO
021050	IF A NOT < 11 AND B < 11 ADD 1 TO F-VOW-CONS ON SIZE	ANAL-TWO
021100	IF A NOT < 11 AND B NOT < 11 ADD 1 TO F-CONS-CONS ON SIZE	ANAL-TWO
021150	IF A ERROR GO TO E-10.	ANAL-TWO
021200	IF A NOT < 11 AND B NOT < 11 ADD 1 TO F-CONS-CONS ON SIZE	ANAL-TWO
021250	IF A ERROR GO TO E-11.	ANAL-TWO
021300	VC-EXIT.	ANAL-TWO
021350	INITIALIZE-RT.	ANAL-TWO
021400	MOVE ZEROS TO VC-TOTAL, A, B, L, M, N, X, Y, F-CONS-CONS,	ANAL-TWO

021450	F-VOW-CONS, F-CONS-VOW, F-VOW-VOW, FOL-FREQ-TOTAL, S,	ANAL-TWO
021500	ILF-TOTAL, CRF-TOTAL, CARD-NK, WS-SEQ-NR.	ANAL-TWO
021550	PERFORM ZEROING-ONE VARYING A FROM 1 BY 1 UNTIL A > 34.	ANAL-TWO
021600	PERFORM ZEROING-TWO VARYING A FROM 1 BY 1 UNTIL A > 34	ANAL-TWO
021650	AFTER B FROM 1 BY 1 UNTIL B > 68.	ANAL-TWO
021700	MOVE LANGUAGE TO LANG-DESIG, OUTPT-FOUR-LANG,	ANAL-TWO
021750	MOVE OUTPT-FOUR-KS-LANG.	ANAL-TWO
021800	MOVE LANG-KEY TO LANG-ID, L-LANG-ID.	ANAL-TWO
021850	CLOSE-OUT-RT.	ANAL-TWO
021900	SUBTRACT 1 FROM ILF (34).	ANAL-TWO
021950	PERFORM IND-TOTAL-CALC VARYING A FROM 1 BY 1 UNTIL A > 34.	ANAL-TWO
022000	PERFORM IND-PROB-CALC VARYING A FROM 1 BY 1 UNTIL A > 34.	ANAL-TWO
022050	PERFORM OUTPT-ONE-PRER.	ANAL-TWO
022100	PERFORM OUTPT-ONE-PRINT VARYING A FROM 1 BY 1 UNTIL A > 34.	ANAL-TWO
022150	MOVE ILF-TOTAL TO ONE-F-ILF-TLT.	ANAL-TWO
022200	WRITE PRINT-LINE FROM OUTPT-ONE-F BEFORE 1.	ANAL-TWO
022250	PERFORM DEPT-TWO-CALC VARYING A FROM 1 BY 1 UNTIL A > 34.	ANAL-TWO
022300	PERFORM OUTPT-TWO-PRER.	ANAL-TWO
022350	PERFORM COL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-TWO
022400	WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.	ANAL-TWO
022450	PERFORM OUTPT-TWO-DUMP-A VARYING B FROM 1 BY 1 UNTIL B > 23.	ANAL-TWO
022500	PERFORM OUTPT-TWO-PRER.	ANAL-TWO
022550	PERFORM COL-TITLES-B VARYING A FROM 14 BY 1 UNTIL A > 26.	ANAL-TWO
022600	WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.	ANAL-TWO
022650	PERFORM OUTPT-TWO-DUMP-B VARYING B FROM 1 BY 1 UNTIL B > 23.	ANAL-TWO
022700	PERFORM OUTPT-TWO-PRER.	ANAL-TWO
022750	PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-TWO
022800	PERFORM COL-TITLES-C VARYING A FROM 27 BY 1 UNTIL A > 34.	ANAL-TWO
022850	WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.	ANAL-TWO
022900	PERFORM OUTPT-TWO-DUMP-C VARYING B FROM 1 BY 1 UNTIL B > 23.	ANAL-TWO
022950	PERFORM OUTPT-TWO-PRER.	ANAL-TWO
023000	PERFORM COL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-TWO
023050	WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.	ANAL-TWO
023100	PERFORM OUTPT-TWO-DUMP-A VARYING B FROM 24 BY 1 UNTIL B > 34.	ANAL-TWO
023150	PERFORM OUTPT-TWO-PRER.	ANAL-TWO
023200	PERFORM COL-TITLES-B VARYING A FROM 14 BY 1 UNTIL A > 26.	ANAL-TWO
023250	WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.	ANAL-TWO
023300	PERFORM OUTPT-TWO-DUMP-B VARYING B FROM 24 BY 1 UNTIL B > 34.	ANAL-TWO
023350	PERFORM OUTPT-TWO-PRER.	ANAL-TWO
023400	PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-TWO
023450	PERFORM COL-TITLES-C VARYING A FROM 27 BY 1 UNTIL A > 34.	ANAL-TWO
023500	WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.	ANAL-TWO
023550	PERFORM OUTPT-TWO-DUMP-C VARYING B FROM 24 BY 1 UNTIL B > 34.	ANAL-TWO
023600	PERFORM OUTPT-TWO-PRER.	ANAL-TWO
023650	PERFORM DEPT-TWO-CALC.	ANAL-TWO
023700	PERFORM COL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13.	ANAL-TWO
023750	WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2.	ANAL-TWO
023800	PERFORM OUTPT-TWO-DUMP-A VARYING B FROM 35 BY 1 UNTIL B > 57.	ANAL-TWO


```

PERFORM OUTPT-THREE-PRP. VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-TWO
WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. UNTIL B > 57. ANAL-TWO
PERFORM OUTPT-TWO-DUMP-B VARYING B FROM 35 BY 1 UNTIL B > 57. ANAL-TWO
PERFORM OUTPT-THREE-PRP. A FROM 1 BY 1 UNTIL A > 13. ANAL-TWO
PERFORM FLUSHING-VARYING A FROM 27 BY 1 UNTIL A > 34. ANAL-TWO
WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. UNTIL B > 57. ANAL-TWO
PERFORM OUTPT-TWO-DUMP-C VARYING B FROM 35 BY 1 UNTIL B > 57. ANAL-TWO
PERFORM OUTPT-THREE-PRP. VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-TWO
WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. UNTIL B > 68. ANAL-TWO
PERFORM OUTPT-TWO-DUMP-A VARYING B FROM 58 BY 1 UNTIL B > 68. ANAL-TWO
PERFORM OUTPT-THREE-PRP. VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-TWO
WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. UNTIL B > 68. ANAL-TWO
PERFORM OUTPT-TWO-DUMP-B VARYING B FROM 58 BY 1 UNTIL B > 68. ANAL-TWO
PERFORM OUTPT-THREE-PRP. A FROM 1 BY 1 UNTIL A > 13. ANAL-TWO
PERFORM FLUSHING-VARYING A FROM 27 BY 1 UNTIL A > 34. ANAL-TWO
WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. UNTIL B > 68. ANAL-TWO
PERFORM OUTPT-TWO-DUMP-C VARYING B FROM 58 BY 1 UNTIL B > 68. ANAL-TWO
PERFORM VC-DUMP. VARYING A FROM 1 BY 1 UNTIL A > 34. ANAL-TWO
PERFORM YULES-PRP. VARYING A FROM 1 BY 1 UNTIL A > 34. ANAL-TWO
COMPUTE YULES-K ROUNDED = (10000 * (S - ILF-TOTAL)) / ANAL-TWO
(ILF-TOTAL ** 2). OUTPT-FOUR-E BEFORE 4. ANAL-TWO
WRITE PRINT-LINE FROM OUTPT-FOUR-F BEFORE 2. ANAL-TWO
WRITE PRINT-LINE FROM OUTPT-FOUR-G BEFORE 2. ANAL-TWO
PERFORM KS-CALC VARYING A FROM 1 BY 1 UNTIL A > 34. ANAL-TWO
MOVE ZEROS TO FOL-FREQ-TOTAL. ANAL-TWO
PERFORM DEPN-TWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 34. ANAL-TWO
PERFORM DEPN-TWO-PROB VARYING B FROM 1 BY 1 UNTIL A > 34. ANAL-TWO
PERFORM DEPN-TWO-PRP. VARYING A FROM 1 BY 1 UNTIL A > 34. ANAL-TWO
PERFORM OUTPT-FIVE-PRP. VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-TWO
WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. B > 34. ANAL-TWO
PERFORM OUTPT-DUMP-D VARYING B FROM 1 BY 1 UNTIL B > 34. ANAL-TWO
PERFORM OUTPT-FIVE-PRP. VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-TWO
WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. B > 34. ANAL-TWO
PERFORM OUTPT-DUMP-E VARYING B FROM 1 BY 1 UNTIL B > 34. ANAL-TWO
PERFORM OUTPT-FIVE-PRP. VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-TWO
PERFORM FLUSHING-VARYING A FROM 27 BY 1 UNTIL A > 34. ANAL-TWO
WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-TWO
PERFORM OUTPT-DUMP-F VARYING B FROM 1 BY 1 UNTIL B > 34. ANAL-TWO

```

```

023850
023900
023950
024000
024050
024100
024150
024200
024250
024300
024350
024400
024450
024500
024550
024600
024650
024700
024750
024800
024850
024900
024950
025000
025050
025100
025150
025200
025250
025300
025350
025400
025450
025500
025550
025600
025650
025700
025750
025800
025850
025900
025950
026000
026050
026100
026150
026200

```



```

026250 MOVE ZEROS TO S. ANAL-TWO
026300 PERFORM YULES-K-PREP VARYING A FROM 1 BY 1 UNTIL A > 34 ANAL-TWO
026350 AFTER B FROM 1 BY 1 UNTIL B > 34. ANAL-TWO
026400 COMPUTE YULES-K-2 ROUNDED = (10000 * (S - FOL-FREQ-TOTAL)) / ANAL-TWO
026450 (FOL-FREQ-TOTAL ** 2). ANAL-TWO
026500 WRITE PRINT-LINE FROM OUTPT-FIVE-B BEFORE 1. ANAL-TWO
026550 MOVE ZEROS TO CRF-TOTAL, WS-SEQ-NR, L. ANAL-TWO
026600 MOVE F, TO TEST, L-TEST. ANAL-TWO
026650 PERFORM CARD-DUMP VARYING A FROM 1 BY 1 UNTIL A > 34 ANAL-TWO
026700 AFTER B FROM 1 BY 1 UNTIL B > 34. ANAL-TWO
026750 MOVE ZEROS TO FOL-FREQ-TOTAL. ANAL-TWO
026800 PERFORM DEPN-TWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 34 ANAL-TWO
026850 AFTER B FROM 35 BY 1 UNTIL B > 68. ANAL-TWO
026900 PERFORM DEPN-TWO-PROB BY 1 VARYING A FROM 1 BY 1 UNTIL A > 34 ANAL-TWO
026950 AFTER B FROM 35 BY 1 UNTIL B > 68. ANAL-TWO
027000 PERFORM OUTPT-SIX-PREP. ANAL-TWO
027050 PERFORM COL-TITLES-A VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-TWO
027100 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-TWO
027150 PERFORM OUTPT-DUMP-D VARYING B FROM 35 BY 1 UNTIL B > 68. ANAL-TWO
027200 PERFORM OUTPT-SIX-PREP. ANAL-TWO
027250 PERFORM COL-TITLES-B VARYING A FROM 14 BY 1 UNTIL A > 26. ANAL-TWO
027300 WRITE PRINT-LINE FROM FIRST-LTP-TITLES BEFORE 2. ANAL-TWO
027350 PERFORM OUTPT-DUMP-E VARYING B FROM 35 BY 1 UNTIL B > 68. ANAL-TWO
027400 PERFORM OUTPT-SIX-PREP. ANAL-TWO
027450 PERFORM FLUSHING VARYING A FROM 1 BY 1 UNTIL A > 13. ANAL-TWO
027500 PERFORM COL-TITLES-C VARYING A FROM 27 BY 1 UNTIL A > 34. ANAL-TWO
027550 WRITE PRINT-LINE FROM FIRST-LTR-TITLES BEFORE 2. ANAL-TWO
027600 PERFORM OUTPT-DUMP-F VARYING B FROM 35 BY 1 UNTIL B > 68. ANAL-TWO
027650 MOVE ZEROS TO S. ANAL-TWO
027700 PERFORM YULES-K-PREP VARYING A FROM 1 BY 1 UNTIL A > 34 ANAL-TWO
027750 AFTER B FROM 35 BY 1 UNTIL B > 68. ANAL-TWO
027800 COMPUTE YULES-K-3 ROUNDED = (10000 * (S - FOL-FREQ-TOTAL)) / ANAL-TWO
027850 (FOL-FREQ-TOTAL ** 2). ANAL-TWO
027900 WRITE PRINT-LINE FROM OUTPT-SIX-B BEFORE 1. ANAL-TWO
027950 MOVE ZEROS TO CRF-TOTAL, WS-SEQ-NR, L. ANAL-TWO
028000 MOVE F, TO TEST, L-TEST. ANAL-TWO
028050 PERFORM CARD-DUMP VARYING A FROM 1 BY 1 UNTIL A > 34 ANAL-TWO
028100 AFTER B FROM 35 BY 1 UNTIL B > 68. ANAL-TWO
028150 IND-TOTAL-CALC. ANAL-TWO
028200 COMPUTE ILF-TOTAL = ILF-TOTAL + ILF (A). ANAL-TWO
028250 IND-PROB-CALC. ANAL-TWO
028300 IF ILF (A) > 0 COMPUTE ILP (A) ROUNDED = ANAL-TWO
028350 ILF (A) / ILF-TOTAL, ELSE MOVE ZEROS TO ILP (A). ANAL-TWO
028400 DEPN-TWO-CALC. ANAL-TWO
028450 MOVE ZEROS TO FOL-FREQ-TOTAL. ANAL-TWO
028500 PERFORM DEPN-TWO-SUM VARYING B FROM 1 BY 1 UNTIL B > 34. ANAL-TWO
028550 PERFORM DEPN-TWO-PROB VARYING B FROM 1 BY 1 UNTIL B > 34. ANAL-TWO
028600 PERFORM DEPN-THREE-CALC. ANAL-TWO

```


035850	MOVE 016450	TO ER-STEP.	ANAL-TWO
035900	GO TO DUMP.		ANAL-TWO
035950	E-2.		ANAL-TWO
036000	MOVE 016850	TO ER-STEP.	ANAL-TWO
036050	GO TO DUMP.		ANAL-TWO
036100	E-3.		ANAL-TWO
036150	MOVE 017300	TO ER-STEP.	ANAL-TWO
036200	GO TO DUMP.		ANAL-TWO
036250	E-4.		ANAL-TWO
036300	MOVE 017550	TO ER-STEP.	ANAL-TWO
036350	GO TO DUMP.		ANAL-TWO
036400	E-5.		ANAL-TWO
036450	MOVE 018200	TO ER-STEP.	ANAL-TWO
036500	GO TO DUMP.		ANAL-TWO
036550	E-6.		ANAL-TWO
036600	MOVE 018850	TO ER-STEP.	ANAL-TWO
036650	GO TO DUMP.		ANAL-TWO
036700	E-7.		ANAL-TWO
036750	MOVE 019800	TO ER-STEP.	ANAL-TWO
036800	GO TO DUMP.		ANAL-TWO
036850	E-8.		ANAL-TWO
036900	MOVE 022450	TO ER-STEP.	ANAL-TWO
036950	GO TO DUMP.		ANAL-TWO
037000	E-9.		ANAL-TWO
037050	MOVE 022500	TO ER-STEP.	ANAL-TWO
037100	GO TO DUMP.		ANAL-TWO
037150	E-10.		ANAL-TWO
037200	MOVE 022600	TO ER-STEP.	ANAL-TWO
037250	GO TO DUMP.		ANAL-TWO
037300	E-11.		ANAL-TWO
037350	MOVE 022700	TO ER-STEP.	ANAL-TWO
037400	GO TO DUMP.		ANAL-TWO
037450	DUMP.		ANAL-TWO
037500	MOVE CARD-NR.	TO ER-CARD.	ANAL-TWO
037550	WRITE PRINT-LINE	FROM ERROR-LINE AFTER 2.	ANAL-TWO
037600	GO TO EOJ.		ANAL-TWO

APPENDIX I. LANGUAGE IDENTIFICATION PROGRAM (V-1)

```

000050 IDENTIFICATION DIVISION.
000100 PROGRAM-ID. LANGUAGE-ID-VERSION-ONE.
000150 AUTHOR. MORTON D. RAU, USN.
000200 REMARKS. THIS PROGRAM IS PART OF MY MASTER OF SCIENCE IN
000250 MANAGEMENT THESIS AND IS DESIGNED TO IDENTIFY SHORT SAMPLE
000300 TEXT OF A LANGUAGE AS EITHER ENGLISH OR SPANISH. THE
000350 DECISION IS BASED ON STATISTICAL CHARACTERISTICS OF LETTERS
000400 WITHIN INDEPENDENT WORDS WITH NO CONSIDERATION GIVEN TO
000450 SPACES.
000500 INPUT DATA MUST APPEAR IN THE FOLLOWING ORDER:
000550 3 ENGLISH A CARDS - ICRF TABLE VALUES X(2)/X(1)
000600 99 ENGLISH B CARDS - DCRF TABLE VALUES X(3)/X(1)
000650 99 ENGLISH C CARDS - DCRF TABLE VALUES
000700 3 SPANISH A CARDS - ICRF TABLE VALUES X(2)/X(1)
000750 99 SPANISH B CARDS - DCRF TABLE VALUES X(3)/X(1)
000800 99 SPANISH C CARDS - DCRF TABLE VALUES X(3)/X(1)
000850 1 HEAD CARD PRECEEDING EACH GROUP OF SAMPLE CARDS.
000900 MAXIMUM CHARACTERS, NOT COUNTING SPACES.
000950 INDIVIDUAL DIVISION.
001000 ENVIRONMENT DIVISION.
001050 CONFIGURATION SECTION.
001100 SOURCE=COMPUTER. IBM-360-67.
001150 OBJECT=COMPUTER. IBM-360-67.
001200 SPECIAL-NAMES.
001250 CO1 IS PAGE-TOP.
001300 INPUT-OUTPUT SECTION.
001350 FILE-CONTROL.
001400 SAMPLE ASSIGN TO UR-S-IN1.
001450 RESULT ASSIGN TO UR-S-OUT1.
001500 DATA DIVISION.
001550 FILE SECTION.
001600 LABEL RECORDS ARE OMITTED
001650 BLOCK CONTAINS 5 RECORDS
001700 DATA RECORD IS SAMPLE-CARD, HEAD-CARD, DATA-CARD.
001750 SAMPLE-FLD.
001800 01 PIC XX.
001850 05 PIC X, OCCURS 71 TIMES.
001900 05 PIC X(7).
001950 05 FILLER
002000 05 HEAD-CARD-TYPE
002050 05 PIC A.
002100 05 NEXT-SAMPLE VALUE X.
002150 05 FILLER
002200

```


002500	05	TEST-NAME	PIC	X(20)-	
002300	05	FILLER	PIC	X(58):	
002350	01	DATA-CARD.			
002400	05	DC-ID	PIC	A.	
002450	05	DC-SEQ	PIC	999.	
002500	05	DC-LANG	PIC	X(10):	
002550	05	CRF-IN	PIC	9V9(5),	OCCURS 11 TIMES.
002600	FD	RESULT			
002650		RECORDS ARE OMITTED			
002700		LABEL CONTAINS 5 RECORDS			
002750		BLOCK CONTAINS 5 RECORDS			
002800		DATA RECORD IS PRINT-LINE	PIC	X(133).	
002850	01	PRINT-LINE			
002900	WORKING-STORAGE SECTION.				
002950	77	A	PIC	99, COMP, SYNC.	
003000	77	B	PIC	99, COMP, SYNC.	
003050	77	N	PIC	99, COMP, SYNC.	
003100	77	S	PIC	9(12), COMP, SYNC.	
003150	77	L	PIC	99, COMP, SYNC.	
003200	77	M	PIC	99, COMP, SYNC.	
003250	77	X	PIC	99, COMP, SYNC.	
003300	77	Y	PIC	99, COMP, SYNC.	
003350	77	VC-TOTAL	PIC	9(4), COMP, SYNC.	
003400	77	F-CONS-CONS	PIC	9(4), COMP, SYNC.	
003450	77	F-VOW-CONS	PIC	9(4), COMP, SYNC.	
003500	77	F-CONS-VOW	PIC	9(4), COMP, SYNC.	
003550	77	F-VOW-VOW	PIC	9(4), COMP, SYNC.	
003600	77	FOL-FREQ-TOTAL	PIC	9(4), COMP, SYNC.	
003650	77	ILF-TOTAL	PIC	9(4), COMP, SYNC.	
003700	77	CARD-NR	PIC	9(3), COMP, SYNC.	
003750	77	K	PIC	9(13)V999.	
003800	77	Z-ENG	PIC	9(5)V999 COMP, SYNC.	
003850	77	Z-SPAN	PIC	9(5)V999 COMP, SYNC.	
003900	77	Z	PIC	99V999 COMP, SYNC.	
003950	77	Z P	PIC	999V9(12) COMP, SYNC.	
004000	77	J	PIC	999V9(6) COMP, SYNC.	
004050	77	Z2	PIC	9V9(5), COMP-3, SYNC.	
004100	77	CRF-TOTAL	PIC	9V9(5), COMP, SYNC.	
004150	77	D	PIC	9V9(5), COMP, SYNC.	
004200	77	ENG-D	PIC	9V9(5), COMP, SYNC.	
004250	77	SPAN-D	PIC	9V9(5), COMP, SYNC.	
004300	77	SAMPLE-SIZE	PIC	9(4), COMP, SYNC.	
004350	77	E	PIC	99V99 COMP, SYNC.	
004400	77	FF12	PIC	999, COMP, SYNC.	
004450	01	FF12 TABLE-VALUES.			
004500	05	FILLER	PIC	X VALUE 'A':	
004550	05	FILLER	PIC	X VALUE 'E':	
004600	05	FILLER	PIC	X VALUE 'I':	

007050	05	ENG-VC-CRF	PIC 9V9(5) OCCURS 4 TIMES.	I DEN-ONE
007100	01	SPAN-VC-CRF-TABLE-VALUES.	PIC 9V9(5) VALUE 0.06163.	I DEN-ONE
007150	05	FILLER	PIC 9V9(5) VALUE 0.43293.	I DEN-ONE
007200	05	FILLER	PIC 9V9(5) VALUE 0.87982.	I DEN-ONE
007250	05	FILLER	PIC 9V9(5) VALUE 1.00000.	I DEN-ONE
007300	05	FILLER	PIC SPAN-VC-CRF REDEFINES.	I DEN-ONE
007350	01	SPAN-VC-CRF	PIC 9V9(5) OCCURS 4 TIMES.	I DEN-ONE
007400	01	IND-FREQ-TABLE.	PIC 9(4) COMP SYNC OCCURS 33 TIMES.	I DEN-ONE
007500	05	ILF	OCCURS 33 TIMES.	I DEN-ONE
007550	01	DEPENDENCY-FREQ-TABLE.	PIC 9(4) COMP SYNC OCCURS 66 TIMES.	I DEN-ONE
007600	05	FST-LTR	OCCURS 33 TIMES.	I DEN-ONE
007650	01	TO FCL-FREQ	OCCURS 33 TIMES.	I DEN-ONE
007700	05	DEPENDENCY-PROB-TABLE.	PIC 9V9(5) COMP SYNC OCCURS 66 TIMES.	I DEN-ONE
007750	05	GIVEN-LTR	OCCURS 33 TIMES.	I DEN-ONE
007800	01	IO FUL-PROB	PIC 9V9(5) COMP SYNC OCCURS 66 TIMES.	I DEN-ONE
007850	05	KS-ALFA-TABLE-VALUES.	PIC V9(6) VALUE .999999.	I DEN-ONE
007900	05	FILLER	PIC V9(6) VALUE .999996.	I DEN-ONE
007950	05	FILLER	PIC V9(6) VALUE .999991.	I DEN-ONE
008000	05	FILLER	PIC V9(6) VALUE .999979.	I DEN-ONE
008050	05	FILLER	PIC V9(6) VALUE .999954.	I DEN-ONE
008100	05	FILLER	PIC V9(6) VALUE .999909.	I DEN-ONE
008150	05	FILLER	PIC V9(6) VALUE .999829.	I DEN-ONE
008200	05	FILLER	PIC V9(6) VALUE .999697.	I DEN-ONE
008250	05	FILLER	PIC V9(6) VALUE .999489.	I DEN-ONE
008300	05	FILLER	PIC V9(6) VALUE .999174.	I DEN-ONE
008350	05	FILLER	PIC V9(6) VALUE .998715.	I DEN-ONE
008400	05	FILLER	PIC V9(6) VALUE .998071.	I DEN-ONE
008450	05	FILLER	PIC V9(6) VALUE .997192.	I DEN-ONE
008500	05	FILLER	PIC V9(6) VALUE .996028.	I DEN-ONE
008550	05	FILLER	PIC V9(6) VALUE .994524.	I DEN-ONE
008600	05	FILLER	PIC V9(6) VALUE .992623.	I DEN-ONE
008650	05	FILLER	PIC V9(6) VALUE .990270.	I DEN-ONE
008700	05	FILLER	PIC V9(6) VALUE .987410.	I DEN-ONE
008750	05	FILLER	PIC V9(6) VALUE .983995.	I DEN-ONE
008800	05	FILLER	PIC V9(6) VALUE .979978.	I DEN-ONE
008850	05	FILLER	PIC V9(6) VALUE .975318.	I DEN-ONE
008900	05	FILLER	PIC V9(6) VALUE .969983.	I DEN-ONE
008950	05	FILLER	PIC V9(6) VALUE .963945.	I DEN-ONE
009000	05	FILLER	PIC V9(6) VALUE .957186.	I DEN-ONE
009050	05	FILLER	PIC V9(6) VALUE .949694.	I DEN-ONE
009100	05	FILLER	PIC V9(6) VALUE .941466.	I DEN-ONE
009150	05	FILLER	PIC V9(6) VALUE .932503.	I DEN-ONE
009200	05	FILLER	PIC V9(6) VALUE .922817.	I DEN-ONE
009250	05	FILLER	PIC V9(6) VALUE .912423.	I DEN-ONE
009300	05	FILLER	PIC V9(6) VALUE .901344.	I DEN-ONE
009350	05	FILLER	PIC V9(6) VALUE .889605.	I DEN-ONE
009400	05	FILLER	PIC V9(6) VALUE .	I DEN-ONE

[illegible]

[illegible]

014250	FILLER	PIC	V9(6)	VALUE	.016378.	IDEN-ONE
014300	FILLER	PIC	V9(6)	VALUE	.015390.	IDEN-ONE
014350	FILLER	PIC	V9(6)	VALUE	.014456.	IDEN-ONE
014400	FILLER	PIC	V9(6)	VALUE	.013574.	IDEN-ONE
014450	FILLER	PIC	V9(6)	VALUE	.012740.	IDEN-ONE
014500	FILLER	PIC	V9(6)	VALUE	.011952.	IDEN-ONE
014550	FILLER	PIC	V9(6)	VALUE	.011209.	IDEN-ONE
014600	FILLER	PIC	V9(6)	VALUE	.010508.	IDEN-ONE
014650	FILLER	PIC	V9(6)	VALUE	.009846.	IDEN-ONE
014700	FILLER	PIC	V9(6)	VALUE	.009223.	IDEN-ONE
014750	FILLER	PIC	V9(6)	VALUE	.008636.	IDEN-ONE
014800	FILLER	PIC	V9(6)	VALUE	.008083.	IDEN-ONE
014850	FILLER	PIC	V9(6)	VALUE	.007562.	IDEN-ONE
014900	FILLER	PIC	V9(6)	VALUE	.007072.	IDEN-ONE
014950	FILLER	PIC	V9(6)	VALUE	.006611.	IDEN-ONE
015000	FILLER	PIC	V9(6)	VALUE	.006177.	IDEN-ONE
015050	FILLER	PIC	V9(6)	VALUE	.005770.	IDEN-ONE
015100	FILLER	PIC	V9(6)	VALUE	.005388.	IDEN-ONE
015150	FILLER	PIC	V9(6)	VALUE	.005028.	IDEN-ONE
015200	FILLER	PIC	V9(6)	VALUE	.004691.	IDEN-ONE
015250	FILLER	PIC	V9(6)	VALUE	.004375.	IDEN-ONE
015300	FILLER	PIC	V9(6)	VALUE	.004078.	IDEN-ONE
015350	FILLER	PIC	V9(6)	VALUE	.003800.	IDEN-ONE
015400	FILLER	PIC	V9(6)	VALUE	.003540.	IDEN-ONE
015450	FILLER	PIC	V9(6)	VALUE	.003296.	IDEN-ONE
015500	FILLER	PIC	V9(6)	VALUE	.003068.	IDEN-ONE
015550	FILLER	PIC	V9(6)	VALUE	.002845.	IDEN-ONE
015600	FILLER	PIC	V9(6)	VALUE	.002645.	IDEN-ONE
015650	KS-ALFA-TABLE	KS-ALFA-TABLE-VALUES.				IDEN-ONE
015700	KS-ALFA	PIC	V9(6)	OCCURS 155 TIMES.		IDEN-ONE
015750	TITLE	PIC	X(38)	VALUE SPACES.		IDEN-ONE
015800	FILLER	PIC	X(38)	VALUE 'IDEN-ONE, ACTUAL SAM		IDEN-ONE
015850	FILLER					IDEN-ONE
015900	PLE LANGUAGE IS:'. .					IDEN-ONE
015950	ANS	PIC	X(20).	VALUE SPACES.		IDEN-ONE
016000	FILLER	PIC	X(37)	VALUE SPACES.		IDEN-ONE
016050	COL-HEAD					IDEN-ONE
016100	FILLER	PIC	X(74)	VALUE SPACES.		IDEN-ONE
016150	FILLER	PIC	X(15)	VALUE 'LANGUAGE.		IDEN-ONE
016200	FILLER	PIC	X(44)	VALUE 'SIGNIFICANCE LEVEL'.		IDEN-ONE
016250	TITLE-UL					IDEN-ONE
016300	FILLER	PIC	X(76)	VALUE SPACES.		IDEN-ONE
016350	FILLER	PIC	X(20)	VALUE ALL '.		IDEN-ONE
016400	FILLER	PIC	X(37)	VALUE SPACES.		IDEN-ONE
016450	COL-HEAD-UL.					IDEN-ONE
016500	FILLER	PIC	X(74)	VALUE SPACES.		IDEN-ONE
016550	FILLER	PIC	X(8)	VALUE ALL '.		IDEN-ONE
016600	FILLER	PIC	X(7)	VALUE SPACES.		IDEN-ONE

019050	05	BANK-SEVEN	PIC X(30).	VALUE SPACES.	IDEN-ONE
019100	05	FILLER	PIC X(30)		IDEN-ONE
019150	01	BANK-FORM.			IDEN-ONE
019200	05	BANK-LANG	PIC X(10).	VALUE SPACES.	IDEN-ONE
019250	05	FILLER	PIC X(10)		IDEN-ONE
019300	05	BANK-EQUAL	PIC X.		IDEN-ONE
019350	05	BANK-ALFA	PIC B9.9(6).		IDEN-ONE
019400	01	ERROR-LINE.			IDEN-ONE
019450	05	FILLER	PIC X(21) VALUE ALL '*'. OVERFLOW ERROR ON:'.		IDEN-ONE
019500	05	FILLER	PIC X(20) VALUE		IDEN-ONE
019550	05	ERROR-ANS	PIC X(20).		IDEN-ONE
019600	05	FILLER	PIC X(72) VALUE ALL '*'. .		IDEN-ONE
019650	01	DATA-FORM.			IDEN-ONE
019700	05	CC	PIC X, OCCURS 73 TIMES.		IDEN-ONE
019750	01	IND-PROB-TABLE.			IDEN-ONE
019800	05	ILP-TABLE.	PIC 9V9(5) COMP SYNC OCCURS 33 TIMES.		IDEN-ONE
019850	01	VC-CRF-TABLE.			IDEN-ONE
019900	05	VC-CRF	PIC 9V9(5) COMP SYNC OCCURS 4 TIMES.		IDEN-ONE
019950	01	O-LINE-EIGHT.			IDEN-ONE
020000	05	FILLER	PIC X.		IDEN-ONE
020050	05	FILLER	PIC X(14) VALUE 'SAMPLE SIZE = '.		IDEN-ONE
020100	05	FILLER-EIGHT	PIC ZZZ9.		IDEN-ONE
020150	05	FILLER	PIC X(29) VALUE ' SINGLE LETTERS CONT		IDEN-ONE
020200	05	AINED ON:'. .			IDEN-ONE
020250	05	O-CARD-NR	PIC ZZ9.		IDEN-ONE
020300	05	FILLER	PIC X(82) VALUE ' CARDS:'. .		IDEN-ONE
020350		PROCEDURE DIVISION.			IDEN-ONE
020400		START-UP.			IDEN-ONE
020450		OPEN INPUT SAMPLE, OUTPUT RESULT.			IDEN-ONE
020500		MOVE 1 TO A.			IDEN-ONE
020550		PERFORM ENG-FILL-1 33 TIMES.			IDEN-ONE
020600		MOVE 1 TO A, B.			IDEN-ONE
020650		PERFORM ENG-FILL-2 1089 TIMES.			IDEN-ONE
020700		MOVE 1 TO A.			IDEN-ONE
020750		MOVE 34 TO B.			IDEN-ONE
020800		PERFORM ENG-FILL-3 1089 TIMES.			IDEN-ONE
020850		MOVE 1 TO A.			IDEN-ONE
020900		PERFORM SPAN-FILL-1 33 TIMES.			IDEN-ONE
021000		MOVE 1 TO A, B.			IDEN-ONE
021050		PERFORM SPAN-FILL-2 1089 TIMES.			IDEN-ONE
021100		MOVE 1 TO A.			IDEN-ONE
021150		MOVE 34 TO B.			IDEN-ONE
021200		PERFORM SPAN-FILL-2 1089 TIMES.			IDEN-ONE
021250		MOVE 1 TO A.			IDEN-ONE
021300		MOVE 34 TO B.			IDEN-ONE
021350		PERFORM SPAN-FILL-2 1089 TIMES.			IDEN-ONE
021400		MOVE 1 TO A.			IDEN-ONE


```

021450 PERFORM SPAN-FILL-3 1089 TIMES.
021500 MOVE ZEROS TO CARD-NR.
021550 BEGIN.
021600 READ SAMPLE AT END GO TO EQJ.
021650 IF NEXT-SAMPLE PERFORM CLOSE-OUT-RT THRU INITIALIZE-RT.
021700 PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.
021750 ADD 1 TO CARD-NR ON SIZE ERROR GO TO ALERT.
021800 MOVE 3 TO N.
021850 DATA-COLL.
021900 IF CC (N) = SPACE GO TO BLANK-COL-STEP.
021950 MOVE 1 TO A.
022000 CK-ONE.
022050 CC (N) = LTR (A) GO TO PROC-TWO.
022100 IF A = 33 GO TO COL-STEP.
022150 ADD 1 TO A.
022200 GO TO CK-ONE.
022250 PROC-TWO.
022300 ADD 1 TO ILF (A) ON SIZE ERROR GO TO ALERT.
022350 COMPUTE M = N + 1.
022400 IF CC (M) = SPACE GO TO COL-STEP.
022450 IF CC (M) = '-' AND M = 73 PERFORM CARD-STEP-ONE.
022500 MOVE 1 TO B.
022550 CK-TWO.
022600 CC (M) = LTR (B) GO TO PROC-THREE.
022650 IF B = 33 GO TO COL-STEP.
022700 ADD 1 TO B.
022750 GO TO CK-TWO.
022800 PROC-THREE.
022850 ADD 1 TO FOL-FREQ (A, B) ON SIZE ERROR GO TO ALERT.
022900 IF A < 11 AND B < 11 ADD 1 TO F-VOW-VOW ON SIZE
022950 IF A ERROR GO TO ALERT.
023000 IF A < 11 AND B NOT < 11 ADD 1 TO F-CONS-VOW ON SIZE
023050 IF A ERROR GO TO ALERT.
023100 IF A NOT < 11 AND B < 11 ADD 1 TO F-VOW-CONS ON SIZE
023150 IF A ERROR GO TO ALERT.
023200 IF A NOT < 11 AND B NOT < 11 ADD 1 TO F-CONS-CONS ON SIZE
023250 IF A ERROR GO TO ALERT.
023300 COMPUTE L = N + 2.
023350 IF CC (L) = SPACE GO TO COL-STEP.
023400 IF CC (L) = '-' AND L = 73 PERFORM CARD-STEP-TWO.
023450 MOVE 1 TO B.
023500 CK-THREE.
023550 CC (L) = LTR (B) PERFORM ADER.
023600 IF CC (L) = LTR (B) GO TO COL-STEP.
023650 IF B = 33 GO TO COL-STEP.
023700 ADD 1 TO B.
023750 GO TO CK-THREE.
023800 COL-STEP.

```



```

0026250 ADD 1 TO L.
0026300 MOVE CRF-IN (L) TO ENG-DCRF (A, B).
0026350 IF B = 33,
0026400 ADD 1 TO A,
0026450 MOVE ZERO TO B.
0026500 ADD 1 TO B.
0026550 IF L = 11 MOVE ZERO TO L.
0026600 ENG-CK-2.
0026650 IF DC-ID NOT = 'B' OR DC-SEQ NOT = CARD-NR
0026700 OR DC-LANG NOT = 'ENGLISH P', GO TO DEFAULT.
0026750 ENG-FILL-3.
0026800 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.
0026850 IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-3.
0026900 ADD 1 TO L.
0026950 MOVE CRF-IN (L) TO ENG-DCRF (A, B).
0027000 IF B = 66
0027050 ADD 1 TO A,
0027100 MOVE 33 TO B.
0027150 ADD 1 TO B.
0027200 IF L = 11 MOVE ZERO TO L.
0027250 ENG-CK-3.
0027300 IF DC-ID NOT = 'C' OR DC-SEQ NOT = CARD-NR
0027350 OR DC-LANG NOT = 'ENGLISH P', GO TO DEFAULT.
0027400 SPAN-FILL-1.
0027450 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.
0027500 IF L = 0 ADD 1 TO CARD-NR PERFORM SPAN-CK-1.
0027550 ADD 1 TO L.
0027600 MOVE CRF-IN (L) TO SPAN-ICRF (A).
0027650 ADD 1 TO A.
0027700 IF L = 11 MOVE ZERO TO L.
0027750 SPAN-CK-1.
0027800 IF DC-ID NOT = 'A' OR DC-SEQ NOT = CARD-NR
0027850 OR DC-LANG NOT = 'SPANISH P', GO TO DEFAULT.
0027900 SPAN-FILL-2.
0027950 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.
0028000 IF L = 0 ADD 1 TO CARD-NR PERFORM SPAN-CK-2.
0028050 ADD 1 TO L.
0028100 MOVE CRF-IN (L) TO SPAN-DCRF (A, B).
0028150 IF B = 33,
0028200 ADD 1 TO A,
0028250 MOVE ZERO TO B.
0028300 ADD 1 TO B.
0028350 IF L = 11 MOVE ZERO TO L.
0028400 SPAN-CK-2.
0028450 IF DC-ID NOT = 'B' OR DC-SEQ NOT = CARD-NR
0028500 OR DC-LANG NOT = 'SPANISH P', GO TO DEFAULT.
0028550 SPAN-FILL-3.
0028600 IF L = 0 READ SAMPLE AT END GO TO DEFAULT.

```




[illegible]

```

PERFORM DEPNTWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 33
AFTER B FROM 1 BY 1 UNTIL B > 33.
PERFORM DEPNTWO-PROB VARYING A FROM 1 BY 1 UNTIL A > 33
AFTER B FROM 1 BY 1 UNTIL B > 33.
MOVE ZEROS TO S.
PERFORM YULES-K-PREP VARYING A FROM 1 BY 1 UNTIL A > 33
AFTER B FROM 1 BY 1 UNTIL B > 33.
COMPUTE K ROUNDED = (10000 * (S - FOL-FREQ-TOTAL)) /
(FOL-FREQ-TOTAL ** 2).
COMPUTE Z-ENG ROUNDED = (K - 87.168) * (CARD-NR ** .5) /
40.207.
COMPUTE Z-SPAN ROUNDED = (K - 104.553) * (CARD-NR ** .5) /
38.094.
PERFORM LANG-CK-K THRU LANG-CK-K-EXIT.
MOVE BANK-FORM TO BANK-TWO.
WRITE PRINT-LINE FROM Q-LINE-TWO BEFORE 2.
PERFORM CR-FOL-PROB-CALC VARYING A FROM 1 BY 1 UNTIL A > 33
AFTER B FROM 1 BY 1 UNTIL B > 33.
MOVE FOL-FREQ-TOTAL TO FFT2.
MOVE ZEROS TO FOL-FREQ-TOTAL.
PERFORM DEPNTWO-SUM VARYING A FROM 1 BY 1 UNTIL A > 33
AFTER B FROM 34 BY 1 UNTIL B > 66.
PERFORM DEPNTWO-PROB VARYING A FROM 1 BY 1 UNTIL A > 33
AFTER B FROM 34 BY 1 UNTIL B > 66.
MOVE ZEROS TO S.
PERFORM YULES-K-PREP VARYING A FROM 1 BY 1 UNTIL A > 33
AFTER B FROM 34 BY 1 UNTIL B > 66.
COMPUTE K ROUNDED = (10000 * (S - FOL-FREQ-TOTAL)) /
(FOL-FREQ-TOTAL ** 2).
COMPUTE Z-ENG ROUNDED = (K - 62.083) * (CARD-NR ** .5) /
42.821.
COMPUTE Z-SPAN ROUNDED = (K - 72.68) * (CARD-NR ** .5) /
35.210.
PERFORM LANG-CK-K THRU LANG-CK-K-EXIT.
MOVE BANK-FORM TO BANK-THREE.
WRITE PRINT-LINE FROM Q-LINE-THREE BEFORE 2.
PERFORM CR-FOL-PROB-CALC VARYING A FROM 1 BY 1 UNTIL A > 33
AFTER B FROM 34 BY 1 UNTIL B > 66.
COMPUTE VC-TOTAL = F-CONS-CCNS + F-VOW-CONS + F-CONS-VOW
+ F-VOW-VOW ON SIZE ERROR GO TO ALERT.
IF F-VOW-VOW > 0
COMPUTE VC-CRF (1) ROUNDED = F-VOW-VOW / VC-TOTAL
ELSE MOVE ZEROS TO VC-CRF (1).
IF F-CONS-VOW > 0
COMPUTE VC-CRF (2) ROUNDED = (F-CONS-VOW / VC-TOTAL)
+ VC-CRF (1)
ELSE MOVE VC-CRF (1) TO VC-CRF (2).
IF F-VOW-CONS > 0

```



```

033450 COMPUTE VC-CRF (3) ROUNDED = (F-VCW-CONS / VC-TOTAL)
033500 + VC-CRF (2)
033550 ELSE MOVE VC-CRF (2) TO VC-CRF (3).
033600 IF F-CONS-CONS > 0
033650 COMPUTE VC-CRF (4) ROUNDED = (F-CONS-CONS / VC-TOTAL)
033700 + VC-CRF (3)
033750 ELSE MOVE VC-CRF (3) TO VC-CRF (4).
033800 PERFORM ENG-VC-DCK VARYING A FROM 1 BY 1 UNTIL A > 4.
033850 PERFORM SPAN-VC-DCK VARYING A FROM 1 BY 1 UNTIL A > 4.
033900 MOVE VC-TOTAL TO SAMPLE-SIZE.
033950 PERFORM LANG-CK-KS.
034000 MOVE BANK-FORM TO BANK-FOUR.
034050 WRITE PRINT-LINE FROM Q-LINE-FOUR BEFORE 2.
034100 MOVE ZEROS TO ENG-D, SPAN-D.
034150 PERFORM ENG-KS-CK1 VARYING A FROM 1 BY 1 UNTIL A > 33.
034200 PERFORM SPAN-KS-CK1 VARYING A FROM 1 BY 1 UNTIL A > 33.
034250 MOVE ILF-TOTAL TO SAMPLE-SIZE.
034300 PERFORM LANG-CK-KS.
034350 MOVE BANK-FORM TO BANK-FIVE.
034400 WRITE PRINT-LINE FROM Q-LINE-FIVE BEFORE 2.
034450 MOVE ZEROS TO ENG-D, SPAN-D.
034500 PERFORM ENG-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 33
034550 AFTER SPAN-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 33
034600 AFTER SPAN-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 33.
034650 MOVE FFT2 TO SAMPLE-SIZE.
034700 PERFORM LANG-CK-KS.
034750 MOVE BANK-FORM TO BANK-SIX.
034800 WRITE PRINT-LINE FROM Q-LINE-SIX BEFORE 2.
034850 MOVE ZEROS TO ENG-D, SPAN-D.
034900 PERFORM ENG-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 33
034950 AFTER SPAN-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 33
035000 AFTER SPAN-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 33
035050 PERFORM SPAN-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 33
035100 AFTER SPAN-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 33.
035150 MOVE FOL-FREQ-TOTAL TO SAMPLE-SIZE.
035200 PERFORM LANG-CK-KS.
035250 MOVE BANK-FORM TO BANK-SEVEN.
035300 WRITE PRINT-LINE FROM Q-LINE-SEVEN BEFORE 2.
035350 WRITE ILF-TOTAL TO BANK-EIGHT.
035400 MOVE CARD-NR TO C-CARD-NR.
035450 WRITE PRINT-LINE FROM Q-LINE-EIGHT AFTER 2.
035500 INITIALIZ-PR.
035550 TEST-NAME TO ANS, ERROR-ANS.
035600 MOVE ZEROS TO A, B, N, S, L, M, X, Y, VC-TOTAL, F-CONS-CONS,
035650 F-VOW-CONS, F-CONS, F-VOW-VOW, FOL-FREQ-TOTAL,
035700 ILF-TOTAL, CARD-NR, K, Z-ENG, Z-SPAN, Z, J, Z2,
035750 CRF-TOTAL, D, ENG-D, SPAN-D, SAMPLE-SIZE, E, I, FFT2.
035800 PERFORM ZEROING-ONE VARYING A FROM 1 BY 1 UNTIL A > 33.

```



```

00358500 PERFORM ZEROING-TWO VARYING A FROM 1 BY 1 UNTIL A > 33
00355900 AFTER B FROM 1 BY 1 UNTIL B > 66.
00355500 READ SAMPLE AT END GO TO DEFAULT.
00360050 LANG-CK-K
00360100 IF Z-ENG = Z-SPAN, TO BANK-LANG,
00360150 MOVE NU DECSN, TO BANK-LANG,
00360200 MOVE Z-ENG TO Z.
00362050 IF Z-ENG > Z-SPAN, TO BANK-LANG,
00362100 MOVE Z-SPAN, TO Z.
00363000 IF Z-ENG < Z-SPAN, TO BANK-LANG,
00363050 MOVE Z-ENG, TO Z.
00364000 IF Z > 4
00365000 MOVE 0.000060 TO BANK-ALFA
00365500 MOVE < TO BANK-EQUAL
00366000 GO TO LANG-CK-K-EXIT.
00366500 COMPUTE P = Z ** 2.
00367000 COMPUTE P = 1 - P ON SIZE ERROR GO TO ALERT.
00367500 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00368000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00368500 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.
00369000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00369500 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00370000 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.
00370500 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00371000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00371500 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.
00372000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00372500 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00373000 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.
00373500 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00374000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00374500 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.
00375000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00375500 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00376000 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.
00376500 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00377000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00377500 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.
00378000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00378500 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00379000 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.
00380000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00380500 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.
00381000 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00381500 COMPUTE J = J + P ON SIZE ERROR GO TO ALERT.
00382000 COMPUTE J = J - P ON SIZE ERROR GO TO ALERT.

```



```

038250 COMPUTE ILF = (P / 34.06451612903) * Z2.
038300 COMPUTE ILF = (P / 36.05060606061) * Z2.
038350 COMPUTE ILF = (P / 38.05714285714) * Z2.
038400 COMPUTE ILF = (P / 40.05405405405) * Z2.
038450 COMPUTE ILF = (P / 42.05128205128) * Z2.
038500 COMPUTE ILF = (P / 44.04878048780) * Z2.
038550 COMPUTE ILF = (P / 46.04651162791) * Z2.
038600 COMPUTE ILF = (P / 48.04444444444) * Z2.
038650 COMPUTE ILF = (P / 50.042553191) * Z2.
038700 COMPUTE ILF = (P / 52.040816326) * Z2.
038750 COMPUTE ILF = (P / 54.039215686) * Z2.
038800 COMPUTE ILF = (P / 56.037735849) * Z2.
038850 COMPUTE ILF = (P / 58.03636) * Z2.
038900 COMPUTE ILF = (P / 60.03509) * Z2.
038950 COMPUTE ILF = (P / 62.03) * Z2.
039000 COMPUTE ILF = (P / 64.03372) * Z2.
039050 COMPUTE ILF = (P / 66.03245) * Z2.
039100 COMPUTE ILF = (P / 68.03118) * Z2.
039150 COMPUTE ILF = (P / 70.02991) * Z2.
039200 COMPUTE ILF = (P / 72.02864) * Z2.
039250 COMPUTE ILF = (P / 74.02737) * Z2.
039300 COMPUTE ILF = (P / 76.02610) * Z2.
039350 COMPUTE ILF = (P / 78.02483) * Z2.
039400 COMPUTE ILF = (P / 80.02356) * Z2.
039450 COMPUTE ILF = (P / 82.02229) * Z2.
039500 COMPUTE ILF = (P / 84.02102) * Z2.
039550 COMPUTE ILF = (P / 86.01975) * Z2.
039600 COMPUTE ILF = (P / 88.01848) * Z2.
039650 COMPUTE ILF = (P / 90.01721) * Z2.
039700 COMPUTE ILF = (P / 92.01594) * Z2.
039750 COMPUTE ILF = (P / 94.01467) * Z2.
039800 COMPUTE ILF = (P / 96.01340) * Z2.
039850 COMPUTE ILF = (P / 98.01213) * Z2.
039900 COMPUTE ILF = (P / 100.01086) * Z2.
039950 COMPUTE ILF = (P / 102.00959) * Z2.
040000 COMPUTE ILF = (P / 104.00832) * Z2.
040050 COMPUTE ILF = (P / 106.00705) * Z2.
040100 COMPUTE ILF = (P / 108.00578) * Z2.
040150 COMPUTE ILF = (P / 110.00451) * Z2.
040200 COMPUTE ILF = (P / 112.00324) * Z2.
040250 COMPUTE ILF = (P / 114.00197) * Z2.
040300 COMPUTE ILF = (P / 116.00070) * Z2.
040350 COMPUTE ILF = (P / 118.00000) * Z2.
040400 COMPUTE ILF = (P / 120.00000) * Z2.
040450 COMPUTE ILF = (P / 122.00000) * Z2.
040500 COMPUTE ILF = (P / 124.00000) * Z2.
040550 COMPUTE ILF = (P / 126.00000) * Z2.
040600 COMPUTE ILF = (P / 128.00000) * Z2.

```



```

0040650 (FOL-FREQ (A, B) / FOL-FREQ-TOTAL), ELSE MOVE ZEROS
0040700 TO FOL-PROB (A, B).
0040750 YULES-K-PRP.
0040800 COMPUTE S = S + (FOL-FREQ (A, B) ** 2).
0040850 CR-FOL-PROB-CALC.
0040900 COMPUTE CRF-TOTAL = CRF-TOTAL + FOL-PROB (A, B).
0040950 MOVE CRF-TOTAL TO FOL-PROB (A, B).
0041000 ENG-VC-DCK.
0041050 COMPUTE D = ENG-VC-CRF (A) - VC-CRF (A).
0041100 IF D > ENG-D MOVE D TO ENG-D.
0041150 SPAN-VC-DCK.
0041200 COMPUTE D = SPAN-VC-CRF (A) - VC-CRF (A).
0041250 IF D > SPAN-D MOVE D TO SPAN-D.
0041300 LANG-CK-KS.
0041350 IF ENG-D = SPAN-D, TO BANK-LANG,
0041400 MOVE ENG-D TO D.
0041450 IF ENG-D > SPAN-D, TO BANK-LANG,
0041500 MOVE SPAN-D TO D.
0041550 IF ENG-D < SPAN-D, TO BANK-LANG,
0041600 MOVE ENG-D TO D.
0041650 COMPUTE Z2 ROUNDED = SAMPLE-SIZE ** .5.
0041700 IF E < Z2,
0041750 MOVE KS-ALFA (1) TO BANK-ALFA.
0041800 IF E > 1.82, TO BANK-EQUAL,
0041850 MOVE KS-ALFA (155) TO BANK-ALFA.
0041900 IF E > 27 AND E < 1.83,
0041950 COMPUTE T ROUNDED = (100 * E) - 27,
0042000 MOVE KS-ALFA (T) TO BANK-ALFA,
0042050 MOVE SPACES TO BANK-EQUAL.
0042100 ENG-KS-CK1.
0042150 COMPUTE D = ENG-ICRF (A) - ILP (A).
0042200 IF D > ENG-D MOVE D TO ENG-D.
0042250 SPAN-KS-CK1.
0042300 COMPUTE D = SPAN-ICRF (A) - ILP (A).
0042350 IF D > SPAN-D MOVE D TO SPAN-D.
0042400 ENG-KS-CK2.
0042450 COMPUTE D = ENG-DCRF (A, B) - FOL-PROB (A, B).
0042500 IF D > ENG-D MOVE D TO ENG-D.
0042550 SPAN-KS-CK2.
0042600 COMPUTE D = SPAN-DCRF (A, B) - FOL-PROB (A, B).
0042650 IF D > SPAN-D MOVE D TO SPAN-D.
0042700 ZEROING-ONE.
0042750
0042800
0042850
0042900
0042950
0043000

```


043050 MOVE ZEROS TO ILF (A), ILP (A).
043100 ZERGING-TWO.
043150 MOVE ZEROS TO FOL-FREQ (A, B), FOL-PROB (A, B).
IDEN-ONE
IDEN-ONE
IDEN-ONE

APPENDIX J. LANGUAGE IDENTIFICATION PROGRAM (V-2)

```

000050 IDENTIFICATION DIVISION.
000100 PROGRAM-ID. LANGUAGE-ID-VERSION-TWO.
000150 AUTHOR. L. MORRISON D. RAU, USN.
000200 REMARKS. THIS PROGRAM IS PART OF MY MASTER OF SCIENCE IN
000250 MANAGEMENT. THIS THESIS AND IS DESIGNED TO IDENTIFY SHORT SAMPLE
000300 TEXT OF A LANGUAGE AS EITHER ENGLISH OR SPANISH. THE
000350 DECISION IS BASED ON STATISTICAL CHARACTERISTICS OF LETTERS,
000400 THE SPACE RATES AND THE INTERDEPENDENCY OF LETTERS IN
000450 ADJOINING WORDS.
000500 INPUT DATA MUST APPEAR IN THE FOLLOWING ORDER:
000550 4 ENGLISH D CARDS - ICRF TABLE VALUES
000600 106 ENGLISH E CARDS - DCRF TABLE VALUES X(2)/X(1)
000650 106 ENGLISH F CARDS - DCRF TABLE VALUES X(3)/X(1)
000700 4 SPANISH D CAPDS - ICRF TABLE VALUES
000750 106 SPANISH E CARDS - DCRF TABLE VALUES X(2)/X(1)
000800 106 SPANISH F CARDS - DCRF TABLE VALUES X(3)/X(1)
000850 1 HEAD CARD PRECEEDING EACH GROUP OF SAMPLE CARDS.
000900 1 HEAD CARD SAMPLE CARDS PER TEST IS 999 OR 9,999
000950 MAXIMUM CHARACTERS, INCLUDING SPACES.
001000 INDIVIDUAL VISION.
001050 ENVIRONMENT SECTION.
001100 CONFIGURATION. IBM-360-67.
001150 SUBJECT-COMPUTER. IBM-360-67.
001200 SPECIAL-NAMES.
001250 CO1 IS PAGE-TOP.
001300 INPUT-OUTPUT SECTION.
001350 FILE-CONTROL. SAMPLE ASSIGN TO UR-S-IN1.
001400 SELECT RESULT ASSIGN TO UR-S-OUT1.
001450 DATA SECTION.
001500 FILE DIVISION.
001550 FD LABEL RECORDS ARE OMITTED
001600 BLOCK CONTAINS 5 RECORDS
001650 DATA RECORD IS SAMPLE-CARD, HEAD-CARD, DATA-CARD.
001700 01 SAMPLE-CARD.
001750 05 FILLER.
001800 05 SAMPLE-FLD.
001850 05 FILLER.
001900 05 FILLER.
001950 05 FILLER.
002000 05 FILLER.
002050 01 HEAD-CARD-TYPE
002100 05 FILLER.
002150 05 FILLER.
002200 05 FILLER.

```


002250	05	TEST-NAME	PIC	X(20).	IDEN-TWO
002300	05	FILLER	PIC <td>X(58).</td> <td>IDEN-TWO</td>	X(58).	IDEN-TWO
002350	01	DATA-CARD.	PIC <td>X(58).</td> <td>IDEN-TWO</td>	X(58).	IDEN-TWO
002400	05	DC-ID	PIC <td>A</td> <td>IDEN-TWO</td>	A	IDEN-TWO
002450	05	DC-SEQ	PIC <td>999.</td> <td>IDEN-TWO</td>	999.	IDEN-TWO
002500	05	DC-LANG	PIC <td>X(10).</td> <td>IDEN-TWO</td>	X(10).	IDEN-TWO
002550	05	CRF-IN	PIC <td>9V9(5),</td> <td>IDEN-TWO</td>	9V9(5),	IDEN-TWO
002600	FD	RESULT		OCCURS 11 TIMES.	IDEN-TWO
002650		RECORDS ARE OMITTED			IDEN-TWO
002700		BLOCK CONTAINS 5 RECORDS			IDEN-TWO
002750		DATA RECORD IS PRINT-LINE	PIC <td>X(133).</td> <td>IDEN-TWO</td>	X(133).	IDEN-TWO
002800	01	PRINT-LINE	PIC <td>99, COMP, SYNC.</td> <td>IDEN-TWO</td>	99, COMP, SYNC.	IDEN-TWO
002850	WORKING	PRINT-STORAGE SECTION.	PIC <td>99, COMP, SYNC.</td> <td>IDEN-TWO</td>	99, COMP, SYNC.	IDEN-TWO
002900	77	A	PIC <td>99, COMP, SYNC.</td> <td>IDEN-TWO</td>	99, COMP, SYNC.	IDEN-TWO
002950	77	B	PIC <td>99, COMP, SYNC.</td> <td>IDEN-TWO</td>	99, COMP, SYNC.	IDEN-TWO
003000	77	N	PIC <td>9(12), COMP, SYNC.</td> <td>IDEN-TWO</td>	9(12), COMP, SYNC.	IDEN-TWO
003050	77	S	PIC <td>99, COMP, SYNC.</td> <td>IDEN-TWO</td>	99, COMP, SYNC.	IDEN-TWO
003100	77	L	PIC <td>99, COMP, SYNC.</td> <td>IDEN-TWO</td>	99, COMP, SYNC.	IDEN-TWO
003150	77	M	PIC <td>99, COMP, SYNC.</td> <td>IDEN-TWO</td>	99, COMP, SYNC.	IDEN-TWO
003200	77	X	PIC <td>99, COMP, SYNC.</td> <td>IDEN-TWO</td>	99, COMP, SYNC.	IDEN-TWO
003250	77	Y	PIC <td>9(4).</td> <td>IDEN-TWO</td>	9(4).	IDEN-TWO
003300	77	VC-TOTAL	PIC <td>9(4), COMP, SYNC.</td> <td>IDEN-TWO</td>	9(4), COMP, SYNC.	IDEN-TWO
003350	77	F-CONS-CONS	PIC <td>9(4), COMP, SYNC.</td> <td>IDEN-TWO</td>	9(4), COMP, SYNC.	IDEN-TWO
003400	77	F-VOW-CONS	PIC <td>9(4), COMP, SYNC.</td> <td>IDEN-TWO</td>	9(4), COMP, SYNC.	IDEN-TWO
003450	77	F-CONS-VOW	PIC <td>9(4), COMP, SYNC.</td> <td>IDEN-TWO</td>	9(4), COMP, SYNC.	IDEN-TWO
003500	77	F-VOW-VOW	PIC <td>9(4), COMP, SYNC.</td> <td>IDEN-TWO</td>	9(4), COMP, SYNC.	IDEN-TWO
003550	77	FOL-FREQ-TOTAL	PIC <td>9(4), COMP, SYNC.</td> <td>IDEN-TWO</td>	9(4), COMP, SYNC.	IDEN-TWO
003600	77	ILF-TOTAL	PIC <td>9(4), COMP, SYNC.</td> <td>IDEN-TWO</td>	9(4), COMP, SYNC.	IDEN-TWO
003650	77	CARD-NR	PIC <td>9(13), V999.</td> <td>IDEN-TWO</td>	9(13), V999.	IDEN-TWO
003700	77	K	PIC <td>9(13), V999.</td> <td>IDEN-TWO</td>	9(13), V999.	IDEN-TWO
003750	77	Z-ENG	PIC <td>9(5), V999.</td> <td>IDEN-TWO</td>	9(5), V999.	IDEN-TWO
003800	77	Z-SPAN	PIC <td>9(5), V999.</td> <td>IDEN-TWO</td>	9(5), V999.	IDEN-TWO
003850	77	P	PIC <td>99V999 COMP, SYNC.</td> <td>IDEN-TWO</td>	99V999 COMP, SYNC.	IDEN-TWO
003900	77	J	PIC <td>99V99(12) COMP, SYNC.</td> <td>IDEN-TWO</td>	99V99(12) COMP, SYNC.	IDEN-TWO
003950	77	Z2	PIC <td>99V99(6) COMP, SYNC.</td> <td>IDEN-TWO</td>	99V99(6) COMP, SYNC.	IDEN-TWO
004000	77	CRF-TOTAL	PIC <td>9V9(5), COMP, SYNC.</td> <td>IDEN-TWO</td>	9V9(5), COMP, SYNC.	IDEN-TWO
004050	77	D	PIC <td>9V9(5), COMP, SYNC.</td> <td>IDEN-TWO</td>	9V9(5), COMP, SYNC.	IDEN-TWO
004100	77	ENG-D	PIC <td>9V9(5), COMP, SYNC.</td> <td>IDEN-TWO</td>	9V9(5), COMP, SYNC.	IDEN-TWO
004150	77	SPAN-D	PIC <td>9V9(5), COMP, SYNC.</td> <td>IDEN-TWO</td>	9V9(5), COMP, SYNC.	IDEN-TWO
004200	77	SAMPLE-SIZE	PIC <td>9V9(5), COMP, SYNC.</td> <td>IDEN-TWO</td>	9V9(5), COMP, SYNC.	IDEN-TWO
004250	77	E	PIC <td>9(4), COMP, SYNC.</td> <td>IDEN-TWO</td>	9(4), COMP, SYNC.	IDEN-TWO
004300	77	FFT2	PIC <td>99V99 COMP, SYNC.</td> <td>IDEN-TWO</td>	99V99 COMP, SYNC.	IDEN-TWO
004400	77	MATCH	PIC <td>9(4).</td> <td>IDEN-TWO</td>	9(4).	IDEN-TWO
004450	77	LTR-	PIC <td>9.</td> <td>IDEN-TWO</td>	9.	IDEN-TWO
004500	01	TABLE-VALUES.	PIC <td>X VALUE 'A'.</td> <td>IDEN-TWO</td>	X VALUE 'A'.	IDEN-TWO
004550	05	FILLER	PIC <td>X VALUE 'E'.</td> <td>IDEN-TWO</td>	X VALUE 'E'.	IDEN-TWO
004600	05	FILLER	PIC <td></td> <td>IDEN-TWO</td>		IDEN-TWO

0004650	05	RRR	VALUE	I.	0	0
0004700	05	LLR	VALUE	U.	1	1
0004750	05	LLR	VALUE	1.	2	2
0004800	05	LLR	VALUE	2.	3	3
0004850	05	LLR	VALUE	3.	4	4
0004900	05	LLR	VALUE	4.	5	5
0004950	05	LLR	VALUE	5.	6	6
0005000	05	LLR	VALUE	A.	7	7
0005050	05	LLR	VALUE	B.	8	8
0005100	05	LLR	VALUE	C.	9	9
0005150	05	LLR	VALUE	D.	10	10
0005200	05	LLR	VALUE	E.	11	11
0005250	05	LLR	VALUE	F.	12	12
0005300	05	LLR	VALUE	G.	13	13
0005350	05	LLR	VALUE	H.	14	14
0005400	05	LLR	VALUE	I.	15	15
0005450	05	LLR	VALUE	J.	16	16
0005500	05	LLR	VALUE	K.	17	17
0005550	05	LLR	VALUE	L.	18	18
0005600	05	LLR	VALUE	M.	19	19
0005650	05	LLR	VALUE	N.	20	20
0005700	05	LLR	VALUE	O.	21	21
0005750	05	LLR	VALUE	P.	22	22
0005800	05	LLR	VALUE	Q.	23	23
0005850	05	LLR	VALUE	R.	24	24
0005900	05	LLR	VALUE	S.	25	25
0005950	05	LLR	VALUE	T.	26	26
0006000	05	LLR	VALUE	U.	27	27
0006050	05	LLR	VALUE	V.	28	28
0006100	05	LLR	VALUE	W.	29	29
0006150	05	LLR	VALUE	X.	30	30
0006200	05	LLR	VALUE	Y.	31	31
0006250	05	LLR	VALUE	Z.	32	32
0006300	05	LLR	VALUE	0.	33	33
0006350	05	LLR	VALUE	1.	34	34
0006400	05	LLR	VALUE	2.	35	35
0006450	05	LLR	VALUE	3.	36	36
0006500	05	LLR	VALUE	4.	37	37
0006550	05	LLR	VALUE	5.	38	38
0006600	05	LLR	VALUE	6.	39	39
0006650	05	LLR	VALUE	7.	40	40
0006700	05	LLR	VALUE	8.	41	41
0006750	05	LLR	VALUE	9.	42	42
0006800	05	LLR	VALUE	A.	43	43
0006850	05	LLR	VALUE	B.	44	44
0006900	05	LLR	VALUE	C.	45	45
0006950	05	LLR	VALUE	D.	46	46
0007000	05	LLR	VALUE	E.	47	47

01	05	LTR-TABLE	REDEFINES LTR-TABLE-VALUES.	PIC X	33 TIMES.
01	05	ENG-ICRF-TABLE.		PIC 9V9(5)	OCCURS 34 TIMES.
01	05	SPAN-ICRF-TABLE.		PIC 9V9(5)	OCCURS 34 TIMES.
01	05	ENG-DCRF-TABLE.		OCCURS 34 TIMES.	
01	05	ENG-ENG-FST-LTR		PIC 9V9(5)	COMP SYNC OCCURS 68 TIMES.
01	05	SPAN-DCRF-TABLE.		OCCURS 34 TIMES.	
01	05	SPAN-SPAN-FST-LTR		PIC 9V9(5)	COMP SYNC OCCURS 68 TIMES.
01	05	ENG-VC-CRF-TABLE-VALUES.		PIC 9V9(5)	VALUE 0.05118.
	05	FILLER		PIC 9V9(5)	VALUE 0.40611.
	05	FILLER		PIC 9V9(5)	VALUE 0.76278.
	05	FILLER		PIC 9V9(5)	VALUE 1.00000.

[illegible]

014250	5	FFFF	KS-ALFA-TABLE	REDEFINES	PIC X(38)	VALUE	SPACES.	IDEN-TWO	TWO
014300	5	FFFF	KS-ALFA		PIC X(38)	VALUE	IDEN-TWO, ACTUAL SAM	IDEN-TWO	TWO
014350	5	FFFF	FFFF		PIC X(20)	VALUE	SPACES.	IDEN-TWO	TWO
014400	5	FFFF	FFFF		PIC X(37)	VALUE	SPACES.	IDEN-TWO	TWO
014450	5	FFFF	FFFF		PIC X(74)	VALUE	SPACES.	IDEN-TWO	TWO
014500	5	FFFF	FFFF		PIC X(15)	VALUE	LANGUAGE.	IDEN-TWO	TWO
014550	5	FFFF	FFFF		PIC X(44)	VALUE	SIGNIFCANCE LEVEL.	IDEN-TWO	TWO
014600	5	FFFF	FFFF		PIC X(76)	VALUE	SPACES.	IDEN-TWO	TWO
014650	5	FFFF	FFFF		PIC X(20)	VALUE	ALL.	IDEN-TWO	TWO
014700	5	FFFF	FFFF		PIC X(37)	VALUE	SPACES.	IDEN-TWO	TWO
014750	5	FFFF	FFFF		PIC X(74)	VALUE	SPACES.	IDEN-TWO	TWO
014800	5	FFFF	FFFF		PIC X(8)	VALUE	ALL.	IDEN-TWO	TWO
014850	5	FFFF	FFFF					IDEN-TWO	TWO
014900	5	FFFF	FFFF					IDEN-TWO	TWO
014950	5	FFFF	FFFF					IDEN-TWO	TWO
015000	5	FFFF	FFFF					IDEN-TWO	TWO
015050	5	FFFF	FFFF					IDEN-TWO	TWO
015100	5	FFFF	FFFF					IDEN-TWO	TWO
015150	5	FFFF	FFFF					IDEN-TWO	TWO
015200	5	FFFF	FFFF					IDEN-TWO	TWO
015250	5	FFFF	FFFF					IDEN-TWO	TWO
015300	5	FFFF	FFFF					IDEN-TWO	TWO
015350	5	FFFF	FFFF					IDEN-TWO	TWO
015400	5	FFFF	FFFF					IDEN-TWO	TWO
015450	5	FFFF	FFFF					IDEN-TWO	TWO
015500	5	FFFF	FFFF					IDEN-TWO	TWO
015550	5	FFFF	FFFF					IDEN-TWO	TWO
015600	5	FFFF	FFFF					IDEN-TWO	TWO
015650	5	FFFF	FFFF					IDEN-TWO	TWO
015700	5	FFFF	FFFF					IDEN-TWO	TWO
015750	5	FFFF	FFFF					IDEN-TWO	TWO
015800	5	FFFF	FFFF					IDEN-TWO	TWO
015850	5	FFFF	FFFF					IDEN-TWO	TWO
015900	5	FFFF	FFFF					IDEN-TWO	TWO
015950	5	FFFF	FFFF					IDEN-TWO	TWO
016000	5	FFFF	FFFF					IDEN-TWO	TWO
016050	5	FFFF	FFFF					IDEN-TWO	TWO
016100	5	FFFF	FFFF					IDEN-TWO	TWO
016150	5	FFFF	FFFF					IDEN-TWO	TWO
016200	5	FFFF	FFFF					IDEN-TWO	TWO
016250	5	FFFF	FFFF					IDEN-TWO	TWO
016300	5	FFFF	FFFF					IDEN-TWO	TWO
016350	5	FFFF	FFFF					IDEN-TWO	TWO
016400	5	FFFF	FFFF					IDEN-TWO	TWO
016450	5	FFFF	FFFF					IDEN-TWO	TWO
016500	5	FFFF	FFFF					IDEN-TWO	TWO
016550	5	FFFF	FFFF					IDEN-TWO	TWO
016600	5	FFFF	FFFF					IDEN-TWO	TWO

016650	FILLER	PIC X(7)	VALUE SPACES.	IDEN-TWO
016700	FILLER	PIC X(18)	VALUE ALL.	IDEN-TWO
016750	FILLER	PIC X(26)	VALUE SPACES.	IDEN-TWO
016800	01-LINE-ONE.			IDEN-TWO
016850	FILLER	PIC X(24)	VALUE SPACES.	IDEN-TWO
016900	FILLER	PIC X(18)	IDENTIFIED BY YULE.	IDEN-TWO
016950	FILLER	PIC X	QUOTE.	IDEN-TWO
017000	FILLER	PIC X(30)	VALUE 'S K FOR SINGLE LETTER	IDEN-TWO
017050	RS AS:			IDEN-TWO
017100	BANK-ONE	PIC X(30)	VALUE SPACES.	IDEN-TWO
017150	FILLER	PIC X(30)	VALUE SPACES.	IDEN-TWO
017200	01-LINE-TWO.			IDEN-TWO
017250	FILLER	PIC X(24)	VALUE SPACES.	IDEN-TWO
017300	FILLER	PIC X(18)	IDENTIFIED BY YULE.	IDEN-TWO
017350	FILLER	PIC X	QUOTE.	IDEN-TWO
017400	FILLER	PIC X(30)	VALUE 'S K FOR X(2)/X(1) AS	IDEN-TWO
017450	:			IDEN-TWO
017500	BANK-TWO	PIC X(30)	VALUE SPACES.	IDEN-TWO
017550	FILLER	PIC X(30)	VALUE SPACES.	IDEN-TWO
017600	01-LINE-THREE.			IDEN-TWO
017650	FILLER	PIC X(24)	VALUE SPACES.	IDEN-TWO
017700	FILLER	PIC X(18)	IDENTIFIED BY YULE.	IDEN-TWO
017750	FILLER	PIC X	QUOTE.	IDEN-TWO
017800	FILLER	PIC X(30)	VALUE 'S K FOR X(3)/X(1) AS	IDEN-TWO
017850	:			IDEN-TWO
017900	BANK-THREE	PIC X(30)	VALUE SPACES.	IDEN-TWO
017950	FILLER	PIC X(30)	VALUE SPACES.	IDEN-TWO
018000	01-LINE-FOUR.			IDEN-TWO
018050	FILLER	PIC X(24)	VALUE SPACES.	IDEN-TWO
018100	FILLER	PIC X(49)	IDENTIFIED BY K-S TE	IDEN-TWO
018150	ST FOR VDW-CONS ORDER	AS:		IDEN-TWO
018200	FILLER	PIC X(30)	VALUE SPACES.	IDEN-TWO
018250	01-LINE-FIVE.			IDEN-TWO
018300	FILLER	PIC X(24)	VALUE SPACES.	IDEN-TWO
018350	FILLER	PIC X(49)	IDENTIFIED BY K-S TE	IDEN-TWO
018400	FILLER			IDEN-TWO
018450	ST FOR SINGLE LETTERS	AS:		IDEN-TWO
018500	FILLER	PIC X(30)	VALUE SPACES.	IDEN-TWO
018550	01-LINE-SIX.			IDEN-TWO
018600	FILLER	PIC X(24)	VALUE SPACES.	IDEN-TWO
018650	FILLER	PIC X(49)	IDENTIFIED BY K-S TE	IDEN-TWO
018700	FILLER			IDEN-TWO
018750	ST FOR X(2)/X(1) AS:			IDEN-TWO
018800	BANK-SIX	PIC X(30)	VALUE SPACES.	IDEN-TWO
018850	FILLER	PIC X(30)	VALUE SPACES.	IDEN-TWO
018900	01-LINE-SEVEN.			IDEN-TWO
018950	FILLER	PIC X(24)	VALUE SPACES.	IDEN-TWO
019000	FILLER	PIC X(49)	IDENTIFIED BY K-S TE	IDEN-TWO


```

019050-      *ST FOR X(3)/X(1) AS:'.
019100    05 BANK-SEVEN
019150    05 FILLER
019200    01 BANK-FORM.
019250    05 FILLER
019300    05 BANK-LANG
019350    05 FILLER
019400    05 BANK-EQUAL
019450    05 BANK-ALFA
019500    01 ERROR-LINE.
019550    05 FILLER
019600    05 FILLER
019650    05 ERROR-ANS
019700    05 FILLER
019750    01 DATA-FORM.
019800    05 CLPG-B-TABLE.
019850    01 IND-PRCB-TABLE.
019900    05 ILP
019950    01 VC-CRF-TABLE.
020000    05 VC-CRF
020050    01 O-LINE-EIGHT.
020100    05 FILLER
020150    05 FILLER
020200    05 BANK-EIGHT
020250    05 FILLER
020300    01 AINED ON '.
020350    05 O-CARD-NR
020400    05 FILLER
020450    01 PROCEDURE DIVISION.
020500    START-UP.
020550    OPEN INPUT SAMPLE, OUTPUT RESULT.
020600    MOVE 1 TO A.
020650    MOVE 0 TO L, CARD-NR.
020700    PERFORM ENG-FILL-1 33 TIMES.
020750    READ 1 TO CARD-NR.
020800    ADD 1 TO CARD-NR.
020850    PERFORM ENG-CK-1.
020900    MOVE 1 TO A, B.
020950    MOVE 0 TO L, CARD-NR.
021000    PERFORM ENG-FILL-2 1155 TIMES.
021050    READ 1 TO CARD-NR.
021100    ADD 1 TO CARD-NR.
021150    PERFORM ENG-CK-2.
021200    MOVE 1 TO A.
021250    MOVE 35 TO B.
021300    MOVE 0 TO L, CARD-NR.
021350    PERFORM ENG-FILL-3 1155 TIMES.
021400    PERFORM ENG-FILL-3 1155 TIMES.

```



```

0021450 READ SAMPLE AT END GO TO DEFAULT.
0021500 ADD 1 TO CARD-NR.
0021550 PERFORM ENG-CK-3.
0021600 MOVE CRF-IN (1) TO ENG-DCRF (34, 68).
0021650 MOVE 1 TO A.
0021700 MOVE 0 TO L, CARD-NR.
0021750 PERFORM SPAN-FILL-1 33 TIMES.
0021800 READ SAMPLE AT END GO TO DEFAULT.
0021850 ADD 1 TO CARD-NR.
0021900 PERFORM SPAN-CK-1.
0021950 MOVE CRF-IN (1) TO SPAN-ICRF (34).
0022000 MOVE 1 TO A, B, CARD-NR.
0022050 MOVE 0 TO L, CARD-NR.
0022100 PERFORM SPAN-FILL-2 155 TIMES.
0022150 READ SAMPLE AT END GO TO DEFAULT.
0022200 ADD 1 TO CARD-NR.
0022250 PERFORM SPAN-CK-2.
0022300 MOVE CRF-IN (1) TO SPAN-DCRF (34, 34).
0022350 MOVE 1 TO A.
0022400 MOVE 35 TO B.
0022450 MOVE 0 TO L, CARD-NR.
0022500 PERFORM SPAN-FILL-3 155 TIMES.
0022550 READ SAMPLE AT END GO TO DEFAULT.
0022600 ADD 1 TO CARD-NR.
0022650 PERFORM SPAN-CK-3.
0022700 MOVE CRF-IN (1) TO SPAN-DCRF (34, 68).
0022750 MOVE ZEROS TO CARD-NR.
0022800 ALPHA.
0022850 PERFORM READ-A-CARD.
0022900 MOVE 2 TO N.
0022950 BRAVO.
0023000 IF CL (N) = SPACE GO TO CK-ONE.
0023050 PERFORM FIRST-LETTER-CHECK THRU FIRST-EXIT.
0023100 IF MATCH = ZERO GO TO CK-ONE.
0023150 ADD 1 TO ILF (A) ON SIZE ERROR GO TO ALERT.
0023200 COMPUTE M = N + 1.
0023250 IF M = 73 = SHIFT-TWO.
0023300 IF CL (M) = SPACE GO TO CK-TWO.
0023350 PERFORM SECOND-LETTER-CHECK THRU SECOND-EXIT.
0023400 IF MATCH = ZERO GO TO CK-TWO.
0023450 CHARLIE.
0023500 ADD 1 TO FOL-FREQ (A, B) ON SIZE ERROR GO TO ALERT.
0023550 PERFORM VC-TALLEY THRU VC-EXIT.
0023600 COMPUTE L = N + 2.
0023650 IF L = 73 AND CL (73) = '-' PERFORM SHIFT-THREE.
0023700 IF L = 73 AND CL (73) NOT = '-', PERFORM ALT-THREE.
0023750 IF CL (L) = SPACE GO TO CK-THREE.
0023800 PERFORM THIRD-LETTER-CHECK THRU THIRD-EXIT.

```



```

023850 IF MATCH = ZERO GO TO CK-THREE.
023900 DELTA.
023950 COMPUTE B = B + 34.
024000 ADD 1 TO FOL-FREQ (A, B) ON SIZE ERROR GO TO ALERT.
024050 ECHO.
024100 ADD 1 TO N.
024150 GO TO BRAVO.
024200 CK-ONE.
024250 ADD 1 TO ILF (34) ON SIZE ERROR GO TO ALERT.
024300 ADD 1 TO N.
024350 MOVE 34 TO A.
024400 CK-ONE-ALPHA.
024450 IF N = 73 PERFORM SHIFT-ONE.
024500 IF CL (N) = SPACE GO TO CK-ONE-BRAVO.
024550 COMPUTE N = N - 1.
024600 PERFORM SECOND-LETTER-CHECK THRU SECOND-EXIT.
024650 IF MATCH = 1 GO TO CHARLIE, ELSE ADD 1 TO N.
024700 CK-ONE-BRAVO.
024750 ADD 1 TO N.
024800 GO TO CK-ONE-ALPHA.
024850 CK-TWO.
024900 ADD 1 TO FOL-FREQ (A, 34) ON SIZE ERROR GO TO ALERT.
024950 ADD 1 TO N.
025000 CK-TWO-ALPHA.
025050 COMPUTE M = N + 1.
025100 IF M = 73 PERFORM SHIFT-TWO.
025150 IF CL (M) = SPACE GO TO CK-TWO-BRAVO.
025200 COMPUTE N = N - 1.
025250 PERFORM THIRD-LETTER-CHECK THRU THIRD-EXIT.
025300 IF MATCH = 1 GO TO DELTA, ELSE ADD 1 TO N.
025350 CK-TWO-BRAVO.
025400 ADD 1 TO N.
025450 GO TO CK-TWO-ALPHA.
025500 CK-THREE.
025550 ADD 1 TO FOL-FREQ (A, 68) ON SIZE ERROR GO TO ALERT.
025600 GO TO ECHO.
025650 SHIFT-ONE.
025700 PERFORM READ-A-CARD.
025750 MOVE SPACES TO CL (1), CL (2).
025800 MOVE 3 TO N.
025850 SHIFT-TWO.
025900 MOVE CL (72) TO CL (2).
025950 PERFORM READ-A-CARD.
026000 MOVE 2 TO N.
026050 MOVE 3 TO N.
026100 SHIFT-THREE.
026150 MOVE CL (72) TO CL (2).
026200 PERFORM READ-A-CARD.

```


026250	MOVE 1 TO N.	I	DEN	-	TWO
026300	MOVE 3 TO L.	I	DEN	-	TWO
026350	ALT-THREE	I	DEN	-	TWO
026400	CL (72) TO CL (1).	I	DEN	-	TWO
026450	MOVE SPACES TO CL (2).	I	DEN	-	TWO
026500	PERFORM READ-A-CARD.	I	DEN	-	TWO
026550	MOVE 0 TO N.	I	DEN	-	TWO
026600	MOVE 2 TO L.	I	DEN	-	TWO
026650	ENG-FILL-1.	I	DEN	-	TWO
026700	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	I	DEN	-	TWO
026750	IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-1.	I	DEN	-	TWO
026800	ADD 1 TO L.	I	DEN	-	TWO
026850	MOVE CRF-IN (L) TO ENG-ICRF (A).	I	DEN	-	TWO
026900	ADD 1 TO A.	I	DEN	-	TWO
026950	IF L = 11 MOVE ZERO TO L.	I	DEN	-	TWO
027000	ENG-CK-1.	I	DEN	-	TWO
027050	IF DC-ID NOT = 'D' OR DC-SEQ NOT = CARD-NR	I	DEN	-	TWO
027100	OR DC-LANG NOT = 'ENGLISH' P, GO TO DEFAULT.	I	DEN	-	TWO
027150	ENG-FILL-2.	I	DEN	-	TWO
027200	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	I	DEN	-	TWO
027250	IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-2.	I	DEN	-	TWO
027300	ADD 1 TO L.	I	DEN	-	TWO
027350	MOVE CRF-IN (L) TO ENG-DCRF (A, B).	I	DEN	-	TWO
027400	IF B = 34, ADD 1 TO A,	I	DEN	-	TWO
027450	MOVE ZERO TO B.	I	DEN	-	TWO
027500	ADD 1 TO B.	I	DEN	-	TWO
027550	IF L = 11 MOVE ZERO TO L.	I	DEN	-	TWO
027600	ENG-CK-2.	I	DEN	-	TWO
027650	IF DC-ID NOT = 'E' OR DC-SEQ NOT = CARD-NR	I	DEN	-	TWO
027700	OR DC-LANG NOT = 'ENGLISH' P, GO TO DEFAULT.	I	DEN	-	TWO
027750	ENG-FILL-3.	I	DEN	-	TWO
027800	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	I	DEN	-	TWO
027850	IF L = 0 ADD 1 TO CARD-NR PERFORM ENG-CK-3.	I	DEN	-	TWO
027900	ADD 1 TO L.	I	DEN	-	TWO
027950	MOVE CRF-IN (L) TO ENG-DCRF (A, B).	I	DEN	-	TWO
028000	IF B = 68, ADD 1 TO A,	I	DEN	-	TWO
028050	MOVE 34 TO B.	I	DEN	-	TWO
028100	ADD 1 TO B.	I	DEN	-	TWO
028150	IF L = 11 MOVE ZERO TO L.	I	DEN	-	TWO
028200	ENG-CK-3.	I	DEN	-	TWO
028250	IF DC-ID NOT = 'F' OR DC-SEQ NOT = CARD-NR	I	DEN	-	TWO
028300	OR DC-LANG NOT = 'ENGLISH' P, GO TO DEFAULT.	I	DEN	-	TWO
028350	SPAN-FILL-1.	I	DEN	-	TWO
028400	IF L = 0 READ SAMPLE AT END GO TO DEFAULT.	I	DEN	-	TWO
028450	IF L = 0 ADD 1 TO CARD-NR PERFORM SPAN-CK-1.	I	DEN	-	TWO
028500	ADD 1 TO L.	I	DEN	-	TWO
028550		I	DEN	-	TWO
028600		I	DEN	-	TWO


```

035850 MOVE BANK-FORM TO BANK-SIX.
035900 WRITE PRINT-LINE FROM O-LINE-SIX BEFORE 2.
035950 MOVE ZEROS TO ENG-D, SPAN-D.
036000 PERF FORM ENG-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 34
036050 AFTER SPAN-B FROM 35 BY 1 UNTIL B > 68.
036100 PERF FORM SPAN-KS-CK2 VARYING A FROM 1 BY 1 UNTIL A > 34
036150 AFTER SPAN-B FROM 35 BY 1 UNTIL B > 68.
036200 MOVE FOL-FREQ-TOTAL TO SAMPLE-SIZE.
036250 PERF FORM LANG-CK-KS.
036300 MOVE BANK-FORM TO BANK-SEVEN.
036350 WRITE PRINT-LINE FROM O-LINE-SEVEN BEFORE 2.
036400 MOVE ILF-TOTAL TO BANK-EIGHT.
036450 MOVE CARD-NR TO O-CARD-NR.
036500 WRITE PRINT-LINE FROM O-LINE-EIGHT AFTER 2.
036550 INITIALIZE-RT.
036600 MOVE TEST-NAME TO ANS, ERROR-ANS.
036650 MOVE ZEROS TO A, B, N, S, L, M, X, Y, VC-TOTAL, F-CONS-CONS,
036700 F-VOW-CONS, F-CONS-VOW, F-VOW-VOW, FOL-FREQ-TOTAL,
036750 ILF-TOTAL, CARD-NR, K, Z-ENG, Z-SPAN, Z, P, J, Z2, FFT2.
036800 CRF-TOTAL, D, ENG-D, SPAN-D, SAMPLE-SIZE, I, A > 34.
036850 PERF FORM ZEROING-ONE VARYING A FROM 1 BY 1 UNTIL A > 34.
036900 PERF FORM ZEROING-TWO VARYING A FROM 1 BY 1 UNTIL A > 34.
036950 AFTER 8 FROM 1 BY 1 UNTIL B > 68.
037000 LANG-CK-K-ENG
037050 IF Z-ENG = Z-SPAN, TO BANK-LANG,
037100 NO DECSN, TO BANK-LANG,
037150 MOVE Z-ENG TO Z.
037200 IF Z-ENG > Z-SPAN, TO BANK-LANG,
037250 MOVE Z-SPAN TO Z.
037300 IF Z-ENG < Z-SPAN, TO BANK-LANG,
037350 MOVE Z-ENG TO Z.
037400 IF Z > 4
037450 MOVE 0.000060 TO BANK-ALFA
037500 MOVE < TO BANK-EQU
037550 GO TO BANK-CK-K-EXIT.
037600 COMPUTE P ROUNDED = (1 / 6) * Z2.
037650 COMPUTE P ROUNDED = (P / 6.666666666667) * Z2.
037700 COMPUTE P ROUNDED = (P / 8.4) * Z2.
037750 COMPUTE P ROUNDED = (P / 10.28571428571) * Z2.
037800 COMPUTE P ROUNDED = (P / 12.22222222222) * Z2.
037850 COMPUTE P ROUNDED = (P / 12.22222222222) * Z2.
037900 COMPUTE P ROUNDED = (P / 12.22222222222) * Z2.
037950 COMPUTE P ROUNDED = (P / 12.22222222222) * Z2.
038000 COMPUTE P ROUNDED = (P / 12.22222222222) * Z2.
038050 COMPUTE P ROUNDED = (P / 12.22222222222) * Z2.
038100 COMPUTE P ROUNDED = (P / 12.22222222222) * Z2.
038150 COMPUTE P ROUNDED = (P / 12.22222222222) * Z2.
038200 COMPUTE P ROUNDED = (P / 12.22222222222) * Z2.

```



```

0043050 IF E > 1.82; TO BANK-EQUAL.
0043100 MOVE K S-ALFA (155) TO BANK-ALFA.
0043150 IF E > 27 AND E < 1.83, (100 * E) - 27,
0043200 COMPUTE T ROUNDED = (100 * E) - 27,
0043250 MOVE K S-ALFA (T) TO BANK-ALFA,
0043300 MOVE SPACES TO BANK-EQUAL.
0043350
0043400 ENG-KS-CK1.
0043450 COMPUTE D = ENG-ICRF (A) - ILP (A).
0043500 IF D > ENG-D MOVE D TO ENG-D.
0043550 SPAN-KS-CK1.
0043600 COMPUTE D = SPAN-ICRF (A) - ILP (A).
0043650 IF D > SPAN-D MOVE D TO SPAN-D.
0043700 ENG-KS-CK2.
0043750 COMPUTE D = ENG-DCRF (A, B) - FOL-PROB (A, B).
0043800 IF D > ENG-D MOVE D TO ENG-D.
0043850 SPAN-KS-CK2.
0043900 COMPUTE D = SPAN-DCRF (A, B) - FOL-PROB (A, B).
0043950 IF D > SPAN-D MOVE D TO SPAN-D.
0044000 ZEROING-ONE.
0044050 MOVE ZEROS TO ILF (A), ILP (A).
0044100 ZEROING-TWO.
0044150 MOVE ZEROS TO FOL-FREQ (A, B), FOL-PROB (A, B).
0044200 READ-A-CARD.
0044250 READ SAMPLE AT END GO TO EOJ.
0044300 IF NEXT-SAMPLE GO TO FLUSH.
0044350 ADD 1 TO CARD-NR ON SIZE ERROR GO TO ALERT.
0044400 PERFORM DAT-SHIFT VARYING X FROM 1 BY 1 UNTIL X > 71.
0044450 CAT-SHIFT.
0044500 COMPUTE Y = X + 2.
0044550 MOVE CC (X) TO CL (Y).
0044600 FLUSH.
0044650 IF CARD-NR NOT = ZERO PERFORM CLOSE-OUT-RT.
0044700 PERFORM INITIALIZE-RT.
0044750 GO TO ALPHA.
0044800 EOJ.
0044850 PERFORM CLOSE-OUT-RT.
0044900 MOVE SPACES TO PRINT-LINE.
0044950 WRITE PRINT-LINE AFTER PAGE-TOP.
0045000 CLOSE SAMPLE, RESULT.
0045050 STOP RUN.
0045100 FIRST-LETTER-CHECK.
0045150 MOVE 1 TO A.
0045200 MOVE ZERO TO MATCH.
0045250 FLC-ALPHA.
0045300 IF CL (N) = LTR (A) GO TO FLC-BRAVO.
0045350 IF A = 33 GO TO FIRST-EXIT.
0045400 ADD 1 TO A.

```


LIST OF REFERENCES

1. Nakamura, Y., "Identification of Languages with Short Sample Texts: A Linguometric Study," Library and Information Science, No. 9, pp 459-81, 1971.
2. Yule, G. U., The Statistical Study of Literary Vocabulary, Cambridge, 1944.
3. Massey, F. J., jr., "A Note on the Estimation of a Distribution Function by Confidence Limits," Annals of Mathematical Statistics, v. 21, pp 116-19, 1950.
4. Birnbaum, Z. W., "Numerical Tabulation of the Distribution of Kolmogorov's Statistic for Finite Sample Size," Journal of the American Statistical Association, v. 47, pp 425-41, September 1952.
5. Birnbaum, Z. W. and Tingey, F. H., "One-Sided Confidence Contours for Probability Distribution Functions," Annals of Mathematical Statistics, v. 22, pp 592-96, 1951.
6. Miller, L. H., "Table of Percentage Points of Kolmogorov Statistic," Journal of the American Statistical Association, v. 51, pp 111-21, March 1956.
7. Smirnov, N., "Tables for Estimating the Goodness of Fit of Empirical Distributions," Annals of Mathematical Statistics, v. 19, pp 279-81, 1948.
8. Lapin, L. L., Statistics for Modern Business Decisions, Jovanovich, 1973.
9. Kelley, T. L., The Kelley Statistical Tables, p. 3, Macmillan, 1938.
10. Altenbernd, L. and Lewis, L. L., Introduction to Literature: Stories, 2nd ed., Macmillan, 1970.
11. Barrett, L. L., Five Centuries of Spanish Literature, Dodd, Mead, & Co., 1963.
12. Medina, Jose Ramon, Antologia Venezolana, Editorial Gredos, 1962.
13. Instituto Internacional de Literatura Iberoamericana, An Anthology of Spanish American Literature, 2nd ed., Appleton-Century-Crofts, 1946.
14. Stallman, R. W. and Watter, R. E., The Creative Reader, 2nd ed., Ronald Press, 1962.

15. Saturday Evening Post, the, Best Modern Short Stories, Curtis Books, 1965.

INITIAL DISTRIBUTION LIST

	No. Copies
1. Defense Documentation Center Cameron Station Alexandria, Virginia 22314	2
2. Library, Code 0212 Naval Postgraduate School Monterey, California 93940	2
3. Department Chairman, Code 55 Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
4. Assistant Professor V. M. Powers, Code 52Pw Department of Electrical Engineering Naval Postgraduate School Monterey, California 93940	1
5. Associate Professor R. A. Weitzman, Code 55Wz Department of Operations Research and Administrative Sciences Naval Postgraduate School Monterey, California 93940	1
6. Commander Eugene J. Normand, Code 52No Naval Postgraduate School Monterey, California 93940	1
7. Commander Naval Telecommunications Command Naval Telecommunications Command Headquarters 4401 Massachusetts Avenue, N. W. Washington, D. C. 20390	1
8. LCDR Morton David Rau, USN, 435-48-2899/1120 3432 Jefferson Avenue New Orleans, Louisiana 70125	1

155642

Thesis

R2446 Rau

c.1

Language identification
by statistical analysis.

155642

Thesis

R2446 Rau

c.1

Language identification
by statistical analysis.

thesR2446

Language identification by statistical a



3 2768 002 05310 0

DUDLEY KNOX LIBRARY